Identification of dynamic oxygen access channels in 12/15-lipoxygenase

Molecular dynamics simulations and free-energy landscapes

DISSERTATION

zur Erlangung des akademischen Grades
doctor rerum naturalium
(Dr. rer. nat.)
im Fach Biophysik

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät I
Humboldt-Universität zu Berlin

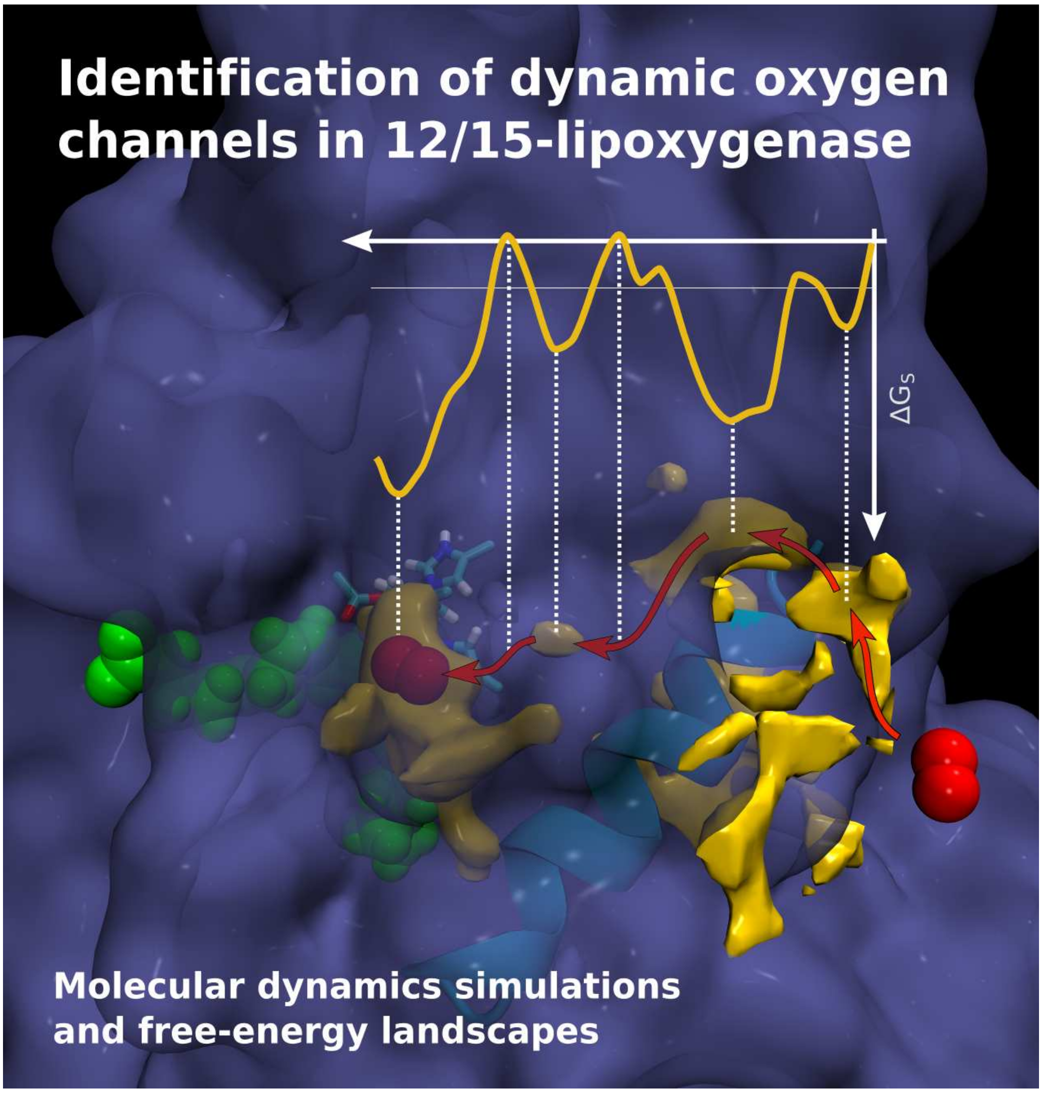
von
Herr Dipl.-Biophys. Jan Saam
geboren am 30.12.1972 in München

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Christoph Markschies

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät I:
Prof. Dr. Christian Limberg

Gutachter:

1. Prof. Dr. Hermann-Georg Holzhütter
2. Prof. Dr. Hartmut Kühn
3. Priv. Doz. Dr. Jörg Grunenberg


Identification of dynamic oxygen channels in 12/15-lipoxygenase
ΔG_S
Molecular dynamics simulations and free-energy landscapes

The cover image

Main oxygen access channel in rabbit 12/15-lipoxygenase. The distribution of oxygen in lipoxygenase is shown in terms of free energy isosurfaces (yellow). Red arrows indicate the energetically most favorable oxygen access route connecting a high affinity region at the protein surface with the catalytic center. Above, the energy profile along this path is projected. The grey line marks the level of the drawn energy isosurface.

Abstract

Cells contain numerous enzymes utilizing molecular oxygen for their reactions. Often, their active sites are buried deeply inside the protein which raises the question whether there are specific access channels guiding oxygen to the site of catalysis. In the present thesis this question is addressed choosing 12/15-lipoxygenase as a typical example for such oxygen dependent enzymes. Lipoxygenases are found in all higher organisms and their products are the precursors of a number of physiological effectors such as inflammation mediators or hormones. They catalyze the position- and stereospecific dioxygenation of fatty acids and lipids to chiral conjugated hydroperoxy compounds. Directed oxygen access to the reaction site through a channel would explain important aspects of the enzyme's stereochemical specificity.

The oxygen distribution within the protein was determined and potential routes for oxygen access were defined. For this purpose an integrated strategy of theoretical and experimental studies including structural modeling, molecular dynamics simulations, site directed mutagenesis, and kinetic measurements was applied.

The limited scope of currently available force fields for biomolecular simulations required the development of force field parameters for the nonheme iron complex in lipoxygenase. The missing parameters were determined based on density functional theory calculations of a simplified model of the coordination complex.

A series of molecular dynamics simulations of the protein in solution was performed. From the trajectories, the 3-dimensional distribution of the free-energy cost for placing oxygen at a certain position could be computed by means of the implicit ligand sampling algorithm. Analyzing energetically favorable paths in the free-energy map led to identification of four oxygen channels in the protein. All channels connect the protein surface with a zone of high oxygen affinity at the active site. This region is localized opposite to the non-heme iron providing a structural explanation for the reaction specificity of this lipoxygenase isoform. Furthermore, it can be seen that the oxygen occupation probability around atom C15 of the arachidonic acid backbone is 7 fold higher than around C11. Thus oxygen insertion at C15 appears to be favored over C11, which is consistent with the positional specificity of the enzyme. However, owing to the high degree of structural flexibility of arachidonic acid the calculations do not exclude oxygen insertion at C11 or at the pro-R side of C15 so that additional mechanisms may contribute to determine the stereochemistry of the oxygenation process.

The catalytically most relevant path can be obstructed by L367F exchange which leads to a strongly increased Michaelis constant for oxygen. This experimentally proven blocking mechanism can, by virtue of molecular dynamics studies of the mutated protein, be explained in detail through a reordering of the hydrogen bonding network of water molecules inside the protein.

As a conclusion, the results of this thesis and the related experiments provide strong evidence that specialized oxygen access channels exist. The main route for oxygen access to the active site of 12/15-lipoxygenase is formed of neighboring, mostly hydrophobic cavities which partition oxygen away from water. These cavities are transiently interconnected due to amino acid side chain flexibility.

Keywords:
lipoxygenase, oxygen channels, molecular dynamics, oxygen diffusion, metalloenzyme, force field, free-energy distribution

Zusammenfassung

Zellen enthalten zahlreiche Enzyme, deren Reaktionen von molekularem Sauerstoff abhängen. Oft sind deren aktive Zentren tief im Inneren des Proteins verborgen, was die Frage nach spezifischen Zugangskanälen, die den Sauerstoff gezielt zum Ort der Katalyse leiten, aufwirft. In der vorliegenden Arbeit wird dies an Hand der 12/15-Lipoxygenase, als ein typisches Beispiel Sauerstoff verbrauchender Enzyme, untersucht. Lipoxygenasen kommen in allen höheren Lebewesen vor und ihre Produkte sind Vorstufen einer Reihe physiologischer Effektoren wie Entzündungsmediatoren oder Hormone. Sie katalysieren die positions- und stereospezifische Dioxygenierung von Fettsäuren und Lipiden zu den entsprechenden Hydroperoxiden. Ein gerichteter Sauerstofftransport zum Reaktionszentrum durch einen Kanal würde helfen, wichtige Aspekte der stereochemischen Spezifität des Enzyms zu erklären.

Um mögliche Routen für den Sauerstoffzugang zu lokalisieren, wurde die Sauerstoffverteilung innerhalb des Proteins bestimmt. Zu diesem Zweck wurden theoretische Untersuchungen basierend auf Molekulardynamik Simulationen eng mit Mutagenese-Experimenten und kinetischen Messungen verzahnt.

Da gängige Kraftfelder für biomolekulare Simulationen nur eine begrenzte Menge unterschiedlicher Komponenten umfassen, war die Entwicklung eigener Kraftfeld-Parameter für den Eisen-Komplex der Lipoxygenase notwendig. Die fehlenden Parameter wurden auf der Basis von Dichtefunktional-Rechnungen eines vereinfachten Modells des Koordinationskomplexes berechnet.

Aus den Trajektorien von Molekulardynamik-Simulationen des Proteins in Lösung konnte die dreidimensionale Verteilung der Freien Enthalpie für Sauerstoff berechnet werden. Die Analyse der günstigsten Pfade in dieser Energielandschaft führte zur Identifikation von vier Sauerstoffkanälen im Protein. Alle Kanäle verbinden die Proteinoberfläche mit einem Gebiet hoher Sauerstoffaffinität am aktiven Zentrum. Diese Region liegt bezüglich des Substrats gegenüber dem Eisenzentrum, wodurch eine strukturelle Erklärung für die Reaktionsspezifität des Enzyms gegeben ist. Weiterhin ist die Aufenthaltswahrscheinlichkeit für Sauerstoff um das C15 Atom der Arachidonsäure herum sieben mal höher als um C11. Sauerstoff scheint also die C15 Insertion zu bevorzugen, was im Einklang mit der Positionsspezifität des Enzyms steht. Dennoch kann man wegen der hohen Konformations-Flexibilität der Arachidonsäure, die auch während der Simulationen beobachtet wurde, eine Insertion an anderen Stellen nicht ausschließen, weshalb auch andere Mechanismen zur beobachteten Stereochemie beitragen könnten.

Der katalytisch bedeutsamste Weg des Sauerstoffs kann durch L367F Austauschmutation blockiert werden, was zu einer stark erhöhten Michaelis-Konstante für Sauerstoff führt. Diese experimentell nachgewiesene Blockade konnte, mit Hilfe entsprechender Molekulardynamik-Simulationen, durch eine Umordnung eines Wasserstoffbrücken-Netzwerks von Wassermolekülen im Protein im Detail erklärt werden.

Die Hauptroute für Sauerstoff zum aktiven Zentrum der 12/15-Lipoxygenase folgt einem Kanal, der aus benachbarten, vorwiegend hydrophoben Hohlräumen besteht. Die Dynamik der Proteinseitenketten sorgt dabei für wiederholte, vorübergehende Verbindung dieser Kammern. Die Ergebnisse dieser Arbeit erlauben den Schluss, dass ähnliche spezialisierte Sauerstoffkanäle auch in anderen sauerstoffabhängigen Enzymen existieren.

Schlagwörter:
Lipoxygenase, Sauerstoffkanäle, Molekulardynamik, Sauerstoffdiffusion, Metalloenzym, Kraftfeld, Freie Enthalpie Verteilung

Widmung

Für meinen Sohn Philipp.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

1.1.1 Oxygen diffusion in proteins

Molecular oxygen plays a role of paramount importance in all aerobic cells. While the largest quantities are reduced by cytochrome c oxidase in the respiratory chain, dioxygen also participates in numerous other enzymatic reactions. Mono- and dioxygenases transfer one and both oxygen atoms from O_2 to a second substrate, respectively. One of the most famous oxygenases is cytochrome P450, which is, in addition to other functions, implicated in steroid synthesis and drug metabolism. Other examples include cyclooxygenases and lipoxygenases. However, except for a few examples little is known about the particular mechanisms how oxygen reaches the reaction sites, which are often buried deep inside the proteins.

Formerly, the hypothesis of unhampered oxygen diffusion through proteins was prevailing. Lakowicz and Weber [1973a] suggested to use the quenching of bound flourophores by molecular oxygen to determine the accessibility of macromolecules for oxygen. Based on tryptophane quenching it was concluded that the oxygen diffusion rate through the protein matrix compares to the one in water [Lakowicz and Weber, 1973b]. Further evidence came from Calhoun et al. [1983] who stated that oxygen diffuses freely through proteins like in a liquid rather than through special channels.

However, this picture is changing since a number of studies have shown the existence of specific oxygen diffusion routes in different proteins. For instance, Riistama et al. [1996] proposed a hydrophobic channel in type bo_3 cytochrome c oxidase that serves as path to the active center deep inside the protein. The observation was based on site-directed mutagenesis studies in which oxygen access was sterically impaired. This is in accordance with results from studies of other isoforms reporting oxygen channels that lead to the catalytic heme group [e.g. Riistama et al., 2000, Koutsoupakis et al., 2003]. Molecular dynamics simulations investigating oxygen diffusion in the protein matrix of cytochrome c oxidase provide further evidence for oxygen migration along a clearly defined channel [Hofacker and Schulten, 1998]. Other examples of enzymes with known oxygen channels include copper amine oxidase [Li et al., 1998] and cholesterol oxidase [Lario et al., 2003].

Lipoxygenases (LOXs) are also interesting candidates for specific oxygen channels. They form a heterogeneous family of nonheme-iron containing fatty acid dioxygenases [Brash, 1999], which are ubiquitously distributed in plants and animals [Feussner and Wasternack, 2002, Kuhn and O'Donnel, 2006]. In mammals LOX-isoforms have been implicated in biosynthesis of inflammatory mediators [Funk, 2001], cell differentiation [van Leyen et al., 1998], pathogenesis of cardiovascular disorders [Kuhn et al., 2005a] and immune regulation [Kuhn and O'Donnel, 2006]. In this thesis oxygen migration routes in rabbit 12/15-LOX are investigated.

The existence of an oxygen access channel in LOXs initially was postulated when the first crystal structure for soybean LOX-1 was solved [Boyington et al., 1993]. An alternative oxygen access path was later suggested [Minor et al., 1996], and recent mutagenesis studies indicate its functionality [Knapp et al., 2001, Knapp and Klinman, 2003]. Inferring from structural comparison that the same oxygen channel also is present in rabbit 12/15-LOX [Gillmor et al., 1997] is intricate due to substantial differences in the relevant parts. Furthermore, the dynamics of oxygen access and the role of a channel for the catalytic specificity remain interesting open questions.

1.1.2 Detecting oxygen channels

Localizing dioxygen in proteins is difficult. One might be able to identify a hydrophobic tunnel in the crystal structure, but how do we know that oxygen is really taking that way? Oxygen is very mobile and usually is not resolved in crystal structures, although mimicking O_2 with xenon, which has a higher electron density, in some cases enabled the detection of well defined oxygen binding sites, e.g. in myoglobin [Tilton et al., 1984]. Other techniques, like tryptophane flourescence quenching, suffer from insufficient resolution. Historically seen, it even lead to wrong conclusions about oxygen access in proteins, as mentioned above. Slowing down oxygen diffusion by steric constriction due to site-directed mutagenesis is useful for testing a hypothetical channel. One problem of this approach is to ascertain that the mutations really influence O_2 diffusion rather than causing other effects such as a modified substrate conformation. As an indirect method it is best used in combination with other approaches.

An alternative consists in the application of computational methods. Molecular dynamics (MD) simulations allow to study the diffusion of individual oxygen molecules in proteins providing a detailed mechanistic interpretation of the protein's interaction with oxygen [e.g. Elber and Karplus, 1990]. However, a single oxygen trajectory represents only one out of many possible paths for oxygen, but actually we are interested in the equilibrium distribution of oxygen in the protein and the most probable routes within that distribution.

The equilibrium distribution of oxygen is determined by the Gibbs free-energy cost $\Delta G_S(\mathrm{O}_2)$ to transfer an oxygen molecule from the solvent to a given position in the protein. If $\Delta G_S(\mathrm{O}_2)$ is negative for a certain region of the protein, then the O_2 concentration in that area is increased. Unfortunately the free-energy distribution cannot be measured directly, and computation based on MD simulations is cumbersome since a large number of different protein conformations have to be taken into account. Recently, implicit ligand sampling (ILS) was introduced [Cohen et al., 2006], a method which allows efficient computation of 3-dimensional free-energy maps for small gas molecules in proteins. The effort is comparably moderate, because only one trajectory in the order of several nanoseconds of the protein without oxygen is needed. An oxygen probe is placed subsequently in many different orientations into the volume elements on a grid. From the resulting samples of interaction energy the free-energy distribution can be computed.

1.2 Objective of this work

The aim of this thesis was to identify potential oxygen access routes in rabbit 12/15-lipoxygenase using computational methods, to suggest supporting experiments for testing the hypothesis and to offer detailed mechanistic explanations for the experimental results. In order to achieve this goal

- the free-energy distribution for oxygen inside the protein was computed using implicit ligand sampling simulations based on MD trajectories of the solvated enzyme with and without bound substrate,
- energetically favorable paths in the 3D free-energy distribution were identified and analyzed,
- critical amino acids were suggested, where exchange mutations might block the hypothetical oxygen channel,
- the mechanism of blocking was elucidated in atomic detail by virtue of MD simulations of a molecular model of the most effective mutant,

- molecular dynamics simulations of explicit oxygen molecules revealing insight into the diffusion dynamics were performed.

An important prerequisit for the MD simulations was to build a molecular model of the solvated enzyme–substrate complex based on the crystallographic structure of 12/15-LOX. The bound inhibitor had to be replaced by arachidonic acid and numerous missing residues had to be completed. Limitations of current force fields for biomolecules required the laborious development of force field parameters for the nonheme-iron complex on the basis of quantum chemical computations.

The results of this work reveal an area of high oxygen affinity at the catalytic center and specific pathways that can be used by oxygen for diffusion into the active site. The directed access provides an explanation for the position- and stereoselectivity of the reaction. Mutation of a residue located at the maximum of the free-energy profile along the main path leads to an increased Michaelis constant for oxygen, i.e. oxygen access is impaired. The channel is blocked due to the formation of a stable hydrogen-bond network involving internal water molecules. Molecular dynamics indicates that oxygen channels in LOX are not just tubes or holes inside the protein matrix, but rather a chain of adjacent, transient association sites being formed and destroyed by the side chain flexibility.

Chapter 2

Background

2.1 Lipoxygenase

2.1.1 Functional characterization

Lipoxygenases catalyze the dioxygenation of fatty acids to the corresponding hydroperoxides. Their main substrates are arachidonic acid and linoleic acid, but the esterified forms of these fatty acids, which are clipped from membrane phospholipids by the enzyme phospholipase A_2, can also be processed. All further mentions of the lipoxygenase substrate will refer to arachidonic acid (hydroperoxy-eicosatetraenoic acid, HPETE), because this fatty acid was used for all studies related to this thesis. LOXs are classified with respect to the position of the carbon atom where oxygen insertion takes place in arachidonic acid (Fig. 2.1). Accordingly, in mammals there are 5-, 8-, 12-, and 15-LOX isoforms forming the products 5-HPETE, 8-HPETE, 12-HPETE, and 15-HPETE, respectively.

The physiological function is best understood for 5-LOX which is predominantly expressed in leukocytes. Its product, 5-HPETE, is transformed in a further enzymatic step into leukotriene A_4 which is the point of origin for the synthesis of other leukotrienes. These substances are mediators of inflammatory reactions like, for instance, phagocyte chemotaxis, increase in vascular permeability or bronchoconstriction [Rådmark et al., 2007, Sharma and Mohammed, 2006]. Leukotrienes are connected with inflammatory diseases like atherosclerosis and rheumatoid arthritis while leukotriene antagonists are used in the treatment of asthma [Bisgaard, 2001].

Much less detail is known about the role of 12/15-LOX, but studies with knockout- and transgenic animals also revealed important roles for 12/15-LOX in inflammatory diseases, including atherosclerosis, cancer, osteoporosis, angiotensin II-dependent hypertension and diabetes. In humans 12/15-LOX is expressed at high levels in reticulocytes (immature red blood cells), eosinophil

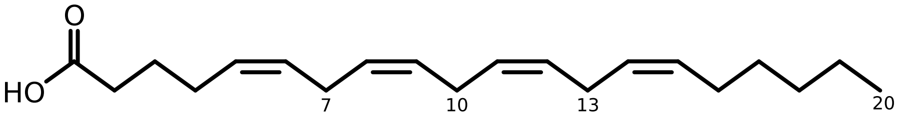


Figure 2.1: Arachidonic acid. Eicosatetraenoic acid is one of the main substrates of LOX. There are bisallylic methylene groups at C7, C10, and C13.

granulocytes (white blood cells of the immune system dealing with parasite infections) and epithelial cells of the bronchia while lower level expression has been reported for numerous other cell types [Kuhn and O'Donnel, 2006]. 12/15-LOX is able to bind to membranes. This association is calcium dependent and is accompanied by an increase of enzyme activity. Oxygenation of membrane lipids by 12/15-LOX seems to play a role in organelle decomposition during red blood cell maturation [Rapoport and Schewe, 1986, Kuhn et al., 1990].

2.1.2 Catalytic Mechanism

The catalytic cycle of LOXs constitutes a redox reaction and involves the formation of radical intermediates. The catalytic center is formed by the nonheme-iron coordination complex which serves as redox equivalent, accepting and donating electrons during the reaction cycle. The beginning of the cycle involves the oxidation of the substrate, where ferric iron (Fe^{3+}) in the active form of the complex accepts an electron from the substrate. Later, in the last catalytic step, the active complex is reconstituted through iron reoxidation. Under physiological conditions the catalytic iron is found in its inactive ferrous form (Fe^{2+}). Prerequisite for the reaction is the activation of the enzyme, i.e. the oxidation of the iron complex. This can be accomplished autocatalytically by hydroperoxy-fatty acids, the LOX reaction products [Yamamoto, 1989]. The catalytic cycle can be subdivided into four steps (see Fig. 2.2):

1) **Hydrogen abstraction:** The first step in catalysis is the abstraction of hydrogen from a bisallylic methylene group (see Fig. 2.1) of the bound substrate by the iron complex. Arachidonic acid has three bisallylic methylene groups, located at C7, C10, and C13. However, the process is position- and stereospecific, i.e. in 15-LOX predominantly the pro-S hydrogen from C13 of arachidonic acid is abstracted. Only a small fraction (7%) of 12-HPETE products are formed due to abstraction of the pro-S hydrogen from C10 [Ivanov et al., 2004]. This selectivity is mainly determined by the location of the substrate in the binding pocket.

   This reaction step exhibits a large temperature-independent isotope effect and a very low enthalpy of activation, which both provide strong evidence

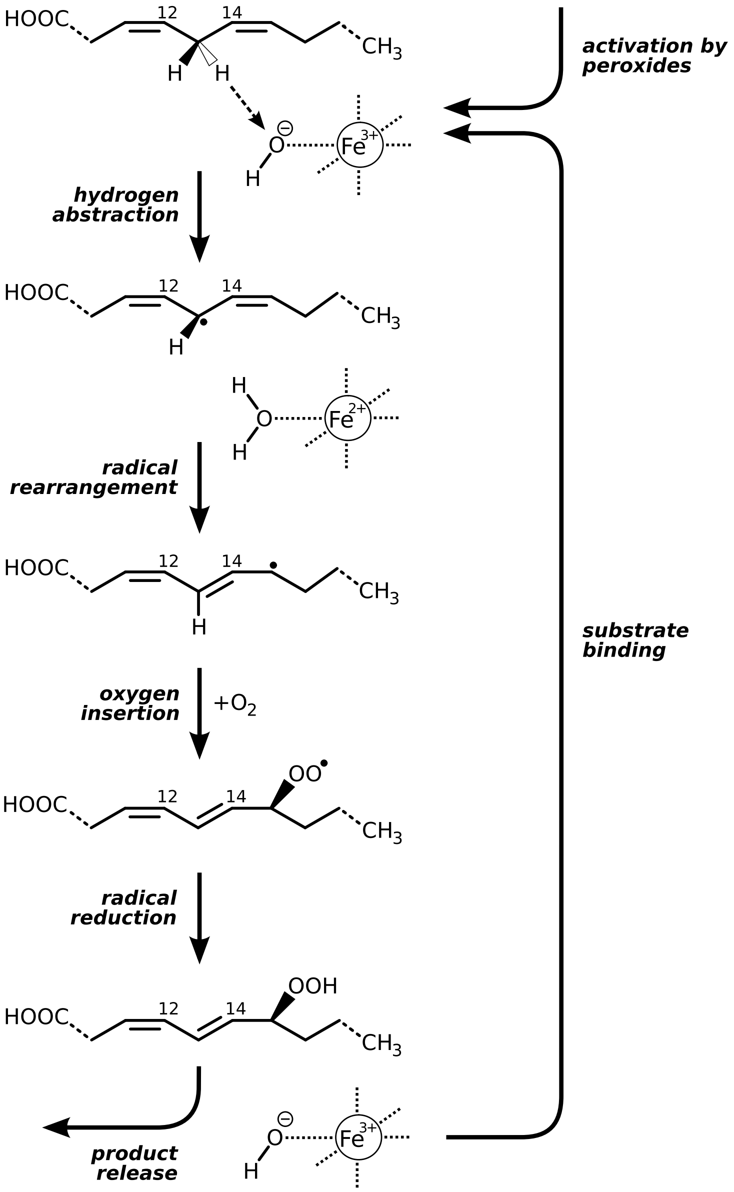


Figure 2.2: Scheme of LOX catalyzed reaction. Carbon numbering refers to arachidonic acid. In order to enter the catalytic cycle ferrous iron has to be oxidized by peroxide species. In the first catalytic step hydrogen is abstracted from the substrate by a tunneling mechanism. The proton reacts with the hydroxide ligand of the iron complex while the electron reduces the ferric iron center. Second, the resulting fatty acid radical is rearranged that the unpaired electron is primarily localized at C15, where molecular oxygen is inserted in the third step. Subsequently, the radical is reduced to the corresponding hydroperoxide, while the activated iron is restored. Finally, the product is released and a new substrate can bind.

for a hydrogen tunneling mechanism [Knapp and Klinman, 2002, Rickert and Klinman, 1999]. Hydrogen tunnelling is the quantum mechanical transfer of the hydrogen atom through rather than over a barrier. Hydrogen abstraction is, despite the low activation enthalpy, the rate-limiting step of the reaction [Egmond et al., 1973], because the tunneling process may show significant entropic barriers, which are dependent on the cross-sectional terms of the tunneling function [Rickert and Klinman, 1999]. In other words, the rate is determined by the probability of the tunneling event, not by the height of the tunneled barrier.

More precisely, the process is described as proton coupled electron transfer. The transferred electron does not localize on the proton, but tunnels directly to the ferric iron in a concerted proton tunneling-electron tunneling process [Lehnert and Solomon, 2003]. But where does the proton go? The ligand pointing towards the substrate is hydroxide or water for the active or inactive complex respectively [Nelson, 1988, Scarrow et al., 1994, Zhang et al., 1995]. The proton resulting from the transfer reacts with the hydroxide ion to form water.

2) **Radical rearrangement:** Homolytic fission of the C–H bond results in a mesomerically stabilized fatty acid radical (Fig. 2.3). Such pentadienyl radicals are well known to add oxygen to form peroxyl radicals [Nelson and Seitz, 1994, Porter, 1986]. Dioxygen would be expected to react at both the C11 and C15 positions, and on the pro-S and pro-R faces, ultimately generating four regio- and stereoisomeric hydroperoxides [Porter, 1986]. Lipoxygenase, however, almost exclusively produces 15-S-peroxide. The position specificity of this reaction requires rearrangement to an allyl-radical (here a conjugated Z-E-dien system), so that the radical electon is mainly localized at C15 where the oxygen insertion will take place. The electronic rearrangement may be explained by a stabilization of enzyme-bound fatty acid radical at the active site in one of its mesomeric limit structures. It may be accomplished by electrophilic amino acid side chains, which focus the electron density at a well defined position or by space-filling residues that force the fatty acid radical into a steric configuration with a more defined localization of the radical electron [Nelson and Seitz, 1994].

Figure 2.3: Arachidonic acid radical. The unpaired electron is delocalized between C11 and C15.

3) **Oxygen Insertion:** Assuming molecular oxygen reacts with C15 of the fatty acid radical, it is still an open question why the insertion is stereospecific. There is an antarafacial relation between initial hydrogen abstraction and subsequent oxygen insertion [Maas and Brash, 1983, Feussner and Wasternack, 2002], i.e. these reaction steps occur on opposite faces with respect to the substrate. The S/R ratio of the 15-LOX products is 97:3 [Saam et al., 2007] (see also Fig. C.1 in Appendix B). It will be shown in this work that molecular dioxygen is not ubiquitously available at the active site, but targeted oxygen access through defined channels in the protein contributes to the enzymes stereo- and position specificity.

4) **Radical Reduction:** In the final step, the radical peroxyl fatty acid will be reduced to the corresponding anion, which can react with a proton to yield the hydroperoxy-fatty acid product. The radical reduction is accompanied by the reoxidation of ferrous to ferric iron, hence the active form of the catalytic center is restored. The product leaves the binding pocket and binding of the next substrate closes the cycle.

2.2 Molecular Dynamics Simulations

Crystallography provides impressive images of biological macromolecules with atomic resolution, yet they are static pictures. Thinking of the function of these molecules, dynamic features like conformational variation, ligand binding or protein folding come to mind. Our dynamic picture relies on spectroscopical observations that contain little or no information about the molecular geometry. At present we have no technique that provides high resolution in time and space at the same time. It would be desirable to watch the molecule at atomic resolution while it performs its function.

One way to address this problem and to study the dynamic behaviour is computer simulation of classical molecular dynamics (MD). Pioneering in the field Alder and Wainwright [1957, 1959] modeled molecules as hard spheres interacting like billiard balls, Stillinger and Rahman [1974] carried out the first simulation of liquid water and in 1977 the first simulation of a protein, the bovine pancreatic trypsin inhibitor (BPTI), appeared [McCammon et al., 1977]. BPTI is a rather small protein (892 atoms) and the simulations at that time were performed *in vacuo*, but the availability of vast computer power nowadays allows for simulations of much larger systems, e.g. proteins in their native environment like water or membrane.

Classical molecular dynamics solves Newton's equations of motion for a system of N atoms interacting according to a potential energy U

$$m_i \frac{\partial^2 r_i}{\partial t^2} = \frac{\partial U}{\partial r_i} = F_i, \tag{2.1}$$

where m_i is the mass and r_i is the position of atom i. The right-hand side denotes the negative gradient of the potential, i.e. the force F_i acting on the atom.

The equation relates the derivative of the potential energy to the changes in position as a function of time. Thus, provided a model of the potential energy U, the so-called force field, is given, integration of equation 2.1 yields the trajectory of the system. The trajectory provides insight into the system's conformational flexibility as the system explores different accessible parts of the phase space.

A molecular model of a protein system for MD simulations is typically based on the crystallographic structure of the protein complemented by solvent molecules and other possible compounds like ligands or membrane. The initial velocities for all atoms are assigned according to the Maxwell-Boltzmann distribution of the desired system temperature.

The choice of the integration time step Δt is crucial for the quality of the simulation. Generally the fidelity increases with shorter time steps. However, to make best use of computer time, a large time step should be used. The C–H bond stretching mode constitutes with $\sim 10^{14}$ Hz the highest vibrational frequency in biomolecular systems. Splitting this period into 10 segments an integration time step of 1 fs is considered sufficient.

Considering the femtosecond timesteps, a serious limitation of MD simulations becomes obvious: In proteins the transitions of interest in proteins often occur on a time scale of microseconds to minutes at 300 K. That means that even for only a millisecond 10^{11} integration steps are necessary. This is further aggravated by the fact that biomolecular systems easily contain 50,000 or even millions of atoms. Such simulations are not yet feasible with current computer power. However, there are also many processes on a shorter timescale, which can be well covered by MD simulations. The problem of oxygen diffusion required a number of simulations in the range of several nanoseconds, still being computationally demanding, but unproblematic on parallel machines.

2.3 Force Fields

2.3.1 Functional form

A force field is constituted by the functional form of the empirical potential describing the interactions between the atoms and by the parameters used in that function [Ermer, 1976, Hagler, 1985]. There exists a large variety of force fields for different applications. The most popular force fields for biomolecular simulations are CHARMM [MacKerell Jr. et al., 1998], AMBER [Cornell et al., 1995], GROMOS96 [Hermans et al., 1984, van Gunsteren et al., 1996] and OPLS-AA [Jorgensen et al., 1996]. They show considerable differences in their parameterization strategy (see Appendix A) and the resulting parameters, but the functional form of these force fields is very similar. It will be outlined in the following.

Classical force fields use empirical functions to describe the atomic interactions. Atoms are treated as spherically symmetric particles, connected through covalent bonds forming molecules. Every atom experiences a force resulting from its interaction with the rest of the system[1]. The expression for the total energy U_{total} includes contributions from bonded and from nonbonded interactions. The bonded interaction encompasses terms for bond stretching U_{bond}, bond angle bending U_{angle}, bond torsion U_{dihed}, and out-of-plane distortions alias improper torsions U_{impr} (Fig. 2.4). Nonbonded interactions are represented by the van der Waals term U_{vdW} and the Coulomb term U_{Coulomb}.

$$U_{\mathrm{total}} = U_{\mathrm{bond}} + U_{\mathrm{angle}} + U_{\mathrm{dihed}} + U_{\mathrm{impr}} + U_{\mathrm{vdW}} + U_{\mathrm{Coulomb}} \tag{2.2}$$

Bond, angle, and improper energies are approximated in terms of harmonic potentials with respect to bond distance r_i, bond angle Θ_i and improper dihedral angle ω_i

$$U_{\mathrm{bond}} = \sum_{\mathrm{bonds}\ i} k_i^{\mathrm{r}}(r_i - r_{0i})^2 \tag{2.3}$$

$$U_{\mathrm{angle}} = \sum_{\mathrm{angles}\ i} k_i^{\Theta}(\Theta_i - \Theta_{0i})^2 \tag{2.4}$$

$$U_{\mathrm{impr}} = \sum_{\mathrm{imprps}\ i} k_i^{\omega}(\omega_i - \omega_{0i})^2, \tag{2.5}$$

while k_i^r, k_i^Θ and k_i^ω denote the force constants. The parameters r_{0i}, Θ_{0i} and ω_{0i} mark the optimum distance and angles at the energy minimum. Bond torsions are characterized by periodic potentials

$$U_{\mathrm{dihed}} = \sum_{\mathrm{diheds}\ i} \frac{k_i^{\Phi}}{2}(1 + \cos(n\Phi_i + \delta_i)) \tag{2.6}$$

with the dihedral angle Φ_i, the barrier height k_i^Φ, and a reference angle δ_i at which the potential is at maximum. The periodicity n_i counts the number of

---

[1] Exception: While the other mentioned force fields are *all-atom* force fields, GROMOS96 uses an *united-atom* approach where methyl groups are treated as single pseudo-atoms.

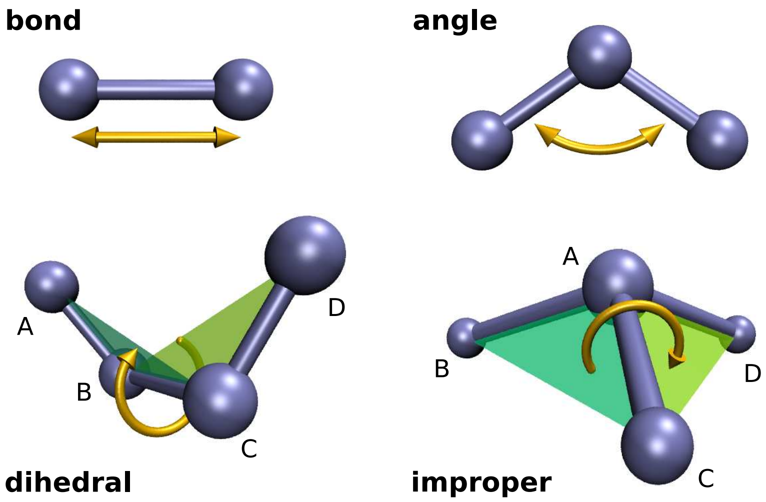


Figure 2.4: Bonded interactions. Bond stretching and bond angle bending should be self-explanatory. Dihedrals measure torsion around bonds defined by the angle between the two planes determined by atoms A–B–C and B–C–D. Improper dihedral torsions can be introduced to control the out-of-plane movement of a central atom bonding to three other atoms. They are defined by the angle between planes A–B–C and A–C–D.

minima for a full rotation of 360° . A sample potential for ethane can be seen in Figure 2.5. In order to account for more complex rotational potential forms it is possible to employ a superposition of multiple sinusoid potentials like Eq. 2.6 with different periodicities, which add up to a Fourier expansion of the potential.

The CHARMM force field knows an additional type of bonded interaction, the Urey-Bradley energy

$$U_{\mathrm{UB}} = \sum_{\mathrm{UB-angles}\ i} k_i^{\mathrm{UB}} (s_i - s_{0i})^2. \tag{2.7}$$

It accounts for angle bending in terms of the distance s_i between atoms A and C in an angle A–B–C. Here, k_i^{UB} is the force constant, and s_{0i} is the equilibrium distance. While all possible bonds, angles, and dihedrals are defined for each molecule, only a limited number of improper and Urey-Bradley terms is introduced in order to optimize the force field's agreement with vibrational data.

The van der Waals (vdW) interaction energy is composed of the attractive dispersion energy due to induced dipole interactions and a repulsive term due to overlapping electron clouds. These phenomena are taken into account by two

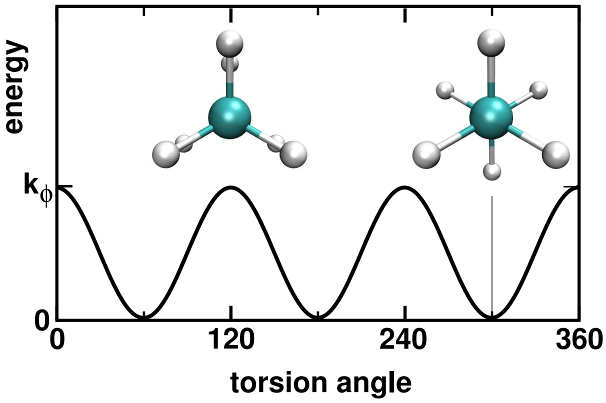


Figure 2.5: Dihedral potential of ethane C–C-bond. Periodicity n=3, phase $\delta = 0$. The barriers are due to the steric interaction of the hydrogen atoms of the two methyl groups. The maxima correspond to eclipsed, minima to the staggered conformation.

terms in the Lennard-Jones potential

$$U_{\mathrm{VDW}} = \sum_{\mathrm{atoms}\ i,j} \epsilon_{ij} \left(\frac{R_{ij}}{r_{ij}}\right)^{12} - 2\left(\frac{R_{ij}}{r_{ij}}\right)^{6} \tag{2.8}$$

that are dependent of the atom distance r_{ij}. The attractive term is proportional to r^{-12}, the repulsive one scales with r^{-6}. Consequently, there exists an optimal distance $R_{ij} = (R_i{+}R_j)/2$ corresponding to a minimum of depth $\epsilon_{ij} = \sqrt{\epsilon_i \cdot \epsilon_j}$ in the potential [2], while at a closer distance repulsion term dominates as illustrated in Figure 2.6.

The second nonbonded term describes the electrostatic interaction of two point charges q_i and q_j

$$U_{\mathrm{Coulomb}} = \sum_{\mathrm{atoms}\ i,j} \frac{q_i q_j}{4\pi\varepsilon\varepsilon_0 r_{ij}}. \tag{2.9}$$

where ε denotes the relative dielectric constant, ε_0 is the permittivity of free space and r_{ij} is the distance between the charges. The atom charges in this expression are an artificial construct since atoms and molecules are not charged, unless they are ions. In molecules, however, the atoms share their valence electrons and the electron density can be shifted due to different electronegativity of the atoms. These different charge densities can be mapped into partial point charges. Note that vdW and Coulomb interactions are usually excluded from computation, if

---

[2] OPLS uses the geometric mean instead for radii and well depth.

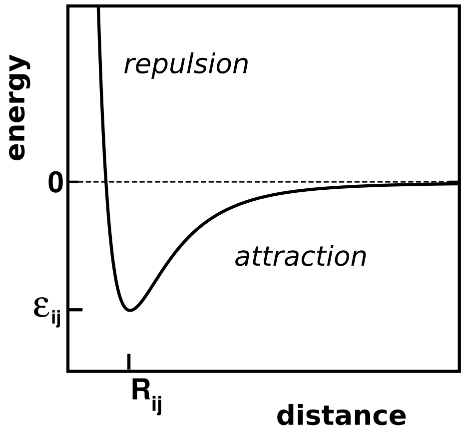


Figure 2.6: Lennard-Jones potential. VdW interaction of atoms i and j where ϵ_{ij} is the potential well depth and optimum atom distance R_{ij}. The repulsive term is $\sim r^{-12}$, the attractive dispersion term is $\sim r^{-6}$. The resulting forces are short ranged and practically vanish after a distance of $\approx$ 1 nm.

the two atoms are separated by less than 3 bonds [3], since these interactions are modeled by the dihedral potential.

2.3.2 Force fields for transition metal complexes

For MD simulations it is crucial to have proper force field parameters for each chemical entity of the system in question. Even though there are approaches like the Universal Force Field (UFF) [Rappé et al., 1992] which can treat all elements in the periodic table and any chemical entity, force fields with a limited parameter set like CHARMM [MacKerell Jr. et al., 1998] and AMBER [Cornell et al., 1995] are far more popular for biomolecular simulations. The reason is that these force fields have been optimized for this class of molecules yielding better results [Mackerell, 2004]. They offer parameters for the most common building blocks including aminoacids, lipids, nucleic acids, and a limited number of cofactors and ligands. Other force fields like Merck Molecular Force Field (MMFF) [Halgren, 1996] or the Generalized Amber Force Field (GAFF) [Wang et al., 2004] were especially designed to cover a great variety of organic ligand molecules.

However, transition metal complexes are not contained in these force fields and cannot be covered by the building block approach because the parameters depend on the metal ion, its coordination number, and the geometric arrangement of the ligands. They have to be derived for each metal complex individually. Even though several force fields for transition metal complexes have been published [Norrby and Brandt, 2001, Banci, 2003], there exists no straightforward, consistent way of parameterizing such complexes. Transfer of the methodology to other systems, especially with high coordination numbers, is troublesome (see

---

[3] For some dihedrals in CHARMM there are special 1–4 interaction terms. In AMBER the nonbonded 1–4 interactions are scaled down by a constant factor

Discussion).

The general importance of these missing parameters immediately becomes clear if one takes into account that transition metal complexes are found as cofactors and redox equivalents in many metalloenzymes. There is for example iron in hydrogenase and nitrogenase, copper in cytochrome oxidase, nickel in urease and zinc in countless different transcription factors as an essential element of their DNA-binding domain.

Lipoxygenase contains a nonheme-iron complex in its catalytic center. For this thesis a new, general methodology to determine CHARMM compliant parameters for transition metal complexes has been developed and applied to the iron complex in LOX. The resulting parameters have been used for all MD simulations of lipoxygenase of this work.

Chapter 3

Methods

3.1 Structural modeling

3.1.1 Protein chain

For MD simulations a solvated model of the LOX-fatty acid substrate complex was required. Since no X-ray coordinates for such a complex are currently available and crystallographic information for the protein [Gillmor et al., 1997] is incomplete structural modeling was performed. LOXs in general and 12/15-LOX in particular consist of two domains. The large catalytic domain, which comprises some 550 amino acids, is mostly helical and contains the substrate-binding pocket with the catalytic non-heme iron. The N-terminal domain, the function of which has not been clarified, is much smaller (111 amino acids) and comprises several anti-parallel β-sheets (Fig 3.1). Since rabbit 12/15-LOX is known to retain its catalytic functionality without the N-terminal domain [Walther et al., 2002], the β-barrel domain (residues 1–111) was omitted in order to simplify the model.

As mentioned above, numerous atoms of the LOX molecule could not be defined exactly in the X-ray structure. The crystal structure has gaps in two surface loops for residues 210–211 and 601–602. In the model for the MD simulations the appropriate amino acids were introduced into these gaps to complete the protein chain. A problematic region in the X-ray data is the sequence between amino acids 177–188, which has ambiguous electron density due to a nearby 2-fold crystallographic symmetry operator. Comparison with the soybean structure [Minor et al., 1996] is not helpful since this part is not well conserved. However, the structures in the sequence immediately before and after the gap are coaxial helices leading to the assumption that the missing piece is helical, too. Several web-based tools were employed to predict the secondary structure for the missing residues. The prediction results of PSIpred [McGuffin et al., 2000], SOPMA

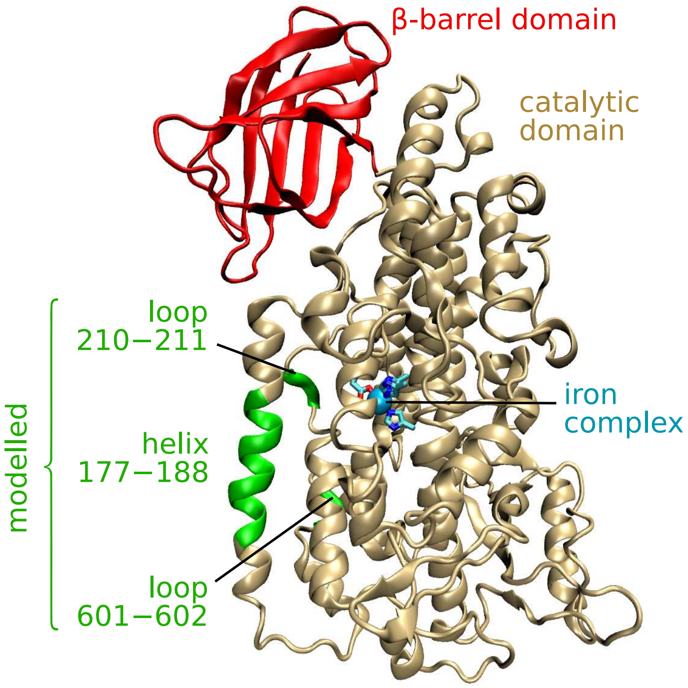


Figure 3.1: Structure of 12/15-LOX. LOX consists of a large catalytic domain (ochre) with the iron complex (blue) at the active center and a β-barrel domain (red), which is not relevant for catalysis. The crystal structure lacks electron density for the surface loop regions 210–211 and 601–602 and for the presumably helical region of residues 177–188. In the structural model for the MD simulations the gaps were completed with the appropriate residues (green) and β-barrel domain was omitted.

[Geourjon and Deléage, 1995], GOR4 [Garnier et al., 1996], and PHDsec [Rost and Sander, 1993] are compared in Fig. 3.2. Even though the results differ for some of the surrounding regions, all secondary structure prediction tools suggest a helical structure for the missing residues. Thus, this region was modeled as a standard helix of the known sequence, which was stretched and twisted in such a way that their ends align well with the two adjacent helices. The modeled helix is located at the protein surface and the side chain of Trp181, which may play a role in membrane binding [Walther et al., 2002], points away from the protein, ideal for membrane interaction.

Histidine side chains were checked for their most likely protonation state, based on their structural environment. They were modeled as neutral residues with protonated N_δ, except for His545, where the proton must be found on N_ϵ, because N_δ is in ideal position to form a complex bond with the iron center. All other titratable amino acids (Lys, Arg, Asp and Glu) in the protein model were

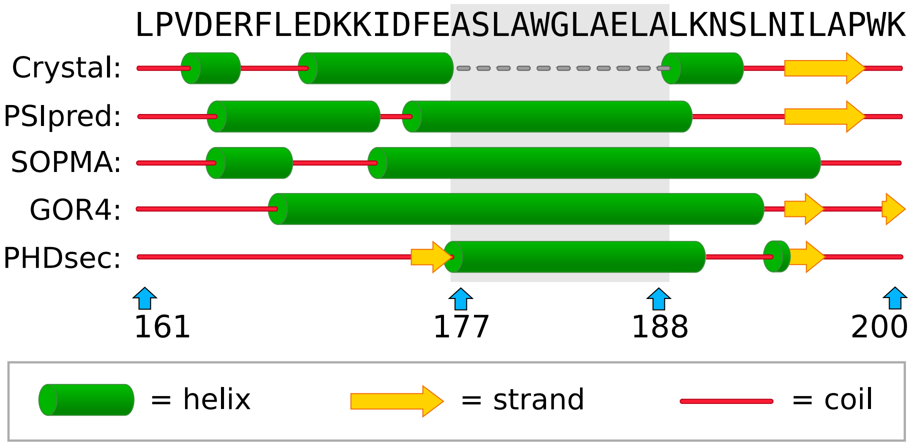


Figure 3.2: Comparison of different secondary structure predictions. In the first line the sequence is given, the second line shows the secondary structure from crystallography, while the other four lines represent the result of the structure prediction from different web-based tools. The region with the gray background corresponds to the gap in the crystal structure between residues 177 and 188.

considered to be charged.

After the research for this thesis had been completed, a reinterpretation of the crystallographic data of rabbit 12/15 LOX was published [Choi et al., 2007], where the coordinates were solved in a different space group. The new structure does not suffer from missing helix atoms, but shows that this region can occur in an open and closed conformation depending on inhibitor binding. It can be shown that the structural model presented herein corresponds well with the closed conformation.

3.1.2 Enzyme–substrate complex

In the crystal structure the non-heme iron is in the ferrous state complexed by five inner shell ligands (4 histidines and a carboxylate oxygen of the C-terminal isoleucine). The sixth ligand position is unoccupied, however, in the structure of the enzyme-inhibitor complex the carboxy group of the inhibitor is found in a weak binding distance of 4.1 Å (Fig. 3.3). According to Lehnert and Solomon [2003] the sixth ligand in the activated Fe(III) coordination complex is a hydroxy ion, which may accept a proton during initial hydrogen abstraction from a bis-allylic methylene of arachidonic acid. For construction of the 12/15-LOX–fatty acid complex the inhibitor was removed and OH^- was added to the iron at an appropriate distance. Next, arachidonic acid was placed into the binding pocket so that the methyl end of the fatty acid was localized close to the triade (F353,

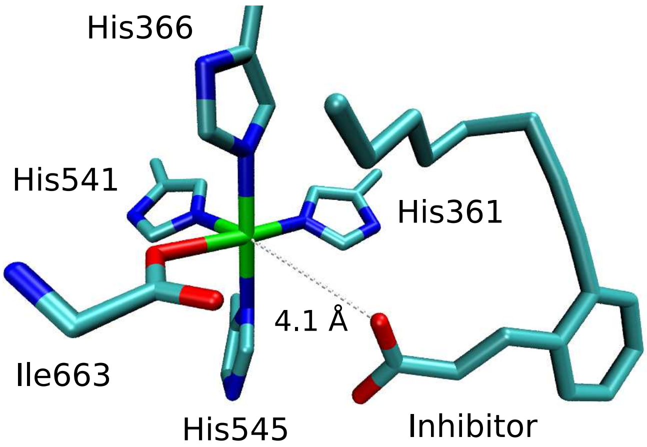


Figure 3.3: Active center with bound inhibitor. The iron complex in the crystal structure of 12/15-LOX is 5-coordinate with an octahedral geometry. The position of the sixth ligand is occupied by an aryl carboxylate inhibitor in the binding pocket, but with a Fe–O distance of 4.1 Å the inhibitor is not part of the inner coordination shell.

I418 and I593) of the sequence determinants of the positional specificity [Kuhn et al., 2005b]. The fatty acid carboxylate was placed in bonding distance to R403 since previous site directed mutagenesis studies suggested ionic interactions between the ionized carboxylate and the positively charged side chain [Gan et al., 1996]. The pro-S hydrogen at carbon 13 of the fatty acid backbone, which is abstracted during the first elementary reaction, was placed in close proximity to the iron so that hydrogen abstraction becomes plausible. The model of the enzyme substrate complex is shown in Fig. 3.4.

3.1.3 Solvation

Internal hydration sites were detected and filled using the computer program DOWSER [Zhang and Hermans, 1996]. This software places water molecules into cavities in the protein, optimizes their orientation and keeps only those molecules with an estimated water–protein interaction energy of -12 kcal/mol or less. The hydrated protein-substrate system was then solvated in water containing chloride and sodium ions at physiological concentration. The number of chloride and sodium ions was chosen to neutralize the net charge of the entire system. The complete model comprised some 44,400 atoms and is shown in Fig. 3.5.

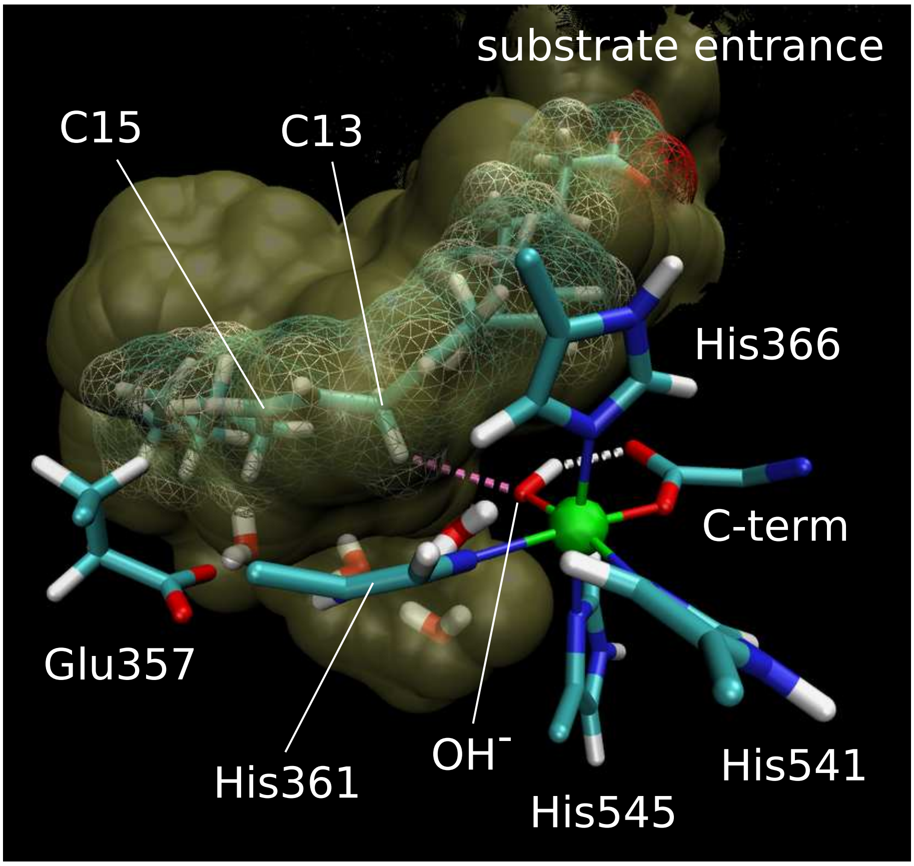


Figure 3.4: Structural model of the 12/15-LOX–arachidonic acid complex. The 3D-model was constructed on the basis of the X-ray coordinates [Gillmor et al., 1997] of the enzyme–inhibitor complex. The central cavity of the enzyme, which constitutes the putative substrate-binding pocket, is indicated in olive. Four histidines and the C-terminal isoleucine form a complex with the catalytic nonheme iron (green). As sixth iron ligand a hydroxyl ion was modeled in. According to quantum chemical calculations, there is a strong hydrogen bridge between the hydroxyl ion and the C-terminus (white dashed line). Besides the bond representation, arachidonic acid is also shown as a space-filling mesh. The main site of oxygen introduction is at C15. The pro-S-hydrogen at C13, which is abstracted during initial hydrogen removal (visualized by the pink dashed line), was placed in close proximity to the iron-bound hydroxyl group. The binding pocket further contains a number of water molecules mainly stabilized by Glu-357 and the OH^--ligand.

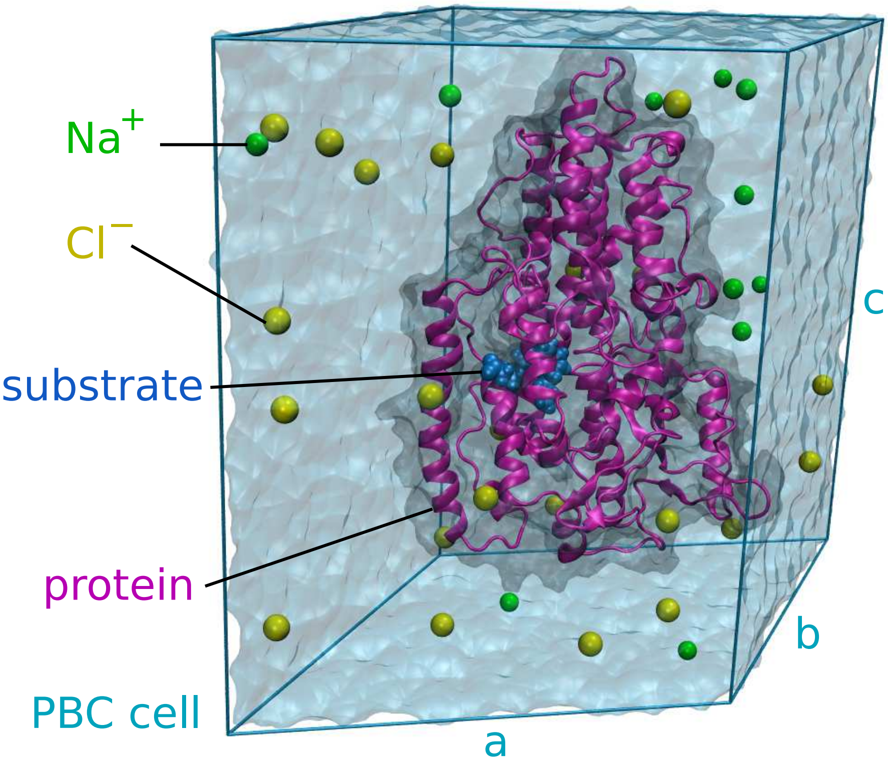


Figure 3.5: Solvated system with periodic boundary conditions. The molecular model used for the MD simulations consists of the protein (purple) with arachidonic acid in the binding pocket (blue). It is solvated in water (light blue box) with a physiological salt concentration. Sodium and chloride ions are shown in green and yellow. The system is simulated in nonorthogonal periodic boundary conditions, the unit-cell is indicated by the blue parallelepiped.

3.2 Force field parameterization

For protein, water, and ions, the CHARMM force field parameters [MacKerell Jr. et al., 1998] were used. For arachidonic acid, the corresponding values were compiled from molecular building blocks of lipids, taken from the same force field database. The parameter set for the 12/15-LOX iron complex was calculated using the program PARATOOL[1] developed by myself. In the following the determination of the parameters for the iron complex is described.

---

[1] PARATOOL is a plugin for the molecular viewer VMD and is distributed with VMD-1.8.5 or greater. Both can be obtained free of charge.

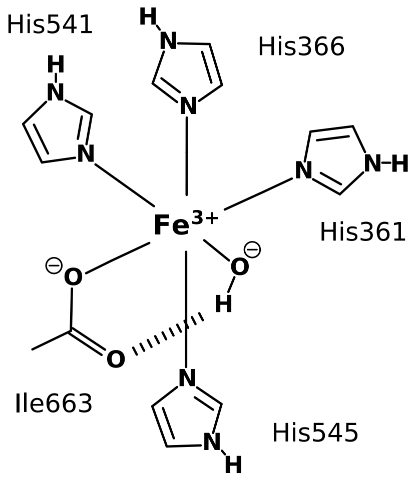


Figure 3.6: Simplified model of the octahedral iron complex in 12/15-LOX in the oxidized state (Fe^{3+}). The ligands are constituted by four imidazoles, the C-terminus of the protein chain and a hydroxy anion pointing towards the substrate. The corresponding residue numbers in the original enzyme are His361, His366, His541, His545 and Ter663. In the reduced state (Fe^{2+}) hydroxide would be replaced by water.

3.2.1 Quantum mechanically derived model of the complex

Parameterization of the iron-complex of LOX requires a simplified, geometry-optimized model of the metal center with ligands from which charges and force constants can be derived. In the model, the four histidines are replaced by imidazoles, and the carboxyl group of the protein chain by acetate. For easier identification, the imidazole ligand models will be called by the name and residue number of their associated amino acid (e.g. His366), the C-terminus at Ile663 will be abbreviated with Ter663. The sixth ligand is hydroxide for the Fe^{3+} or water for the Fe^{2+} configuration (Fig. 3.6).

The exact geometry is known for neither oxidation state. As can be seen in Fig. 3.3, the crystal structure was obtained for inactivated LOX (Fe^{2+}) with an inhibitor bound in the binding pocket [Gillmor et al., 1997]. Although the geometry is octahedral and the inhibitor places a carboxylate moiety in the vicinity of the complex, this group does not participate in the coordination due to the large distance of (~ 4 Å). From EPR studies [e.g. Carroll et al., 1993, Zhang et al., 1995] it is known, that the electronic configuration for both, active and inactive complex, is high spin.

Starting from the crystal structure coordinates (including the manually added hydroxide/water ligand) the molecular geometry of the simplified model was optimized employing density functional theory (DFT) [Hohenberg and Kohn, 1964, Kohn and Sham, 1965, Kleinman, 1984] with the Becke3LYP functional [Becke, 1988] and the 6-31+G* basis set. The calculations were performed for both oxidation states. All DFT calculations were performed with Gaussian03 [Frisch et al., 2003].

3.2.2 Parameterization Scheme

A set of force field parameters was generated for the iron-complex in the oxidized form. Assuming that parameters for the lone ligands are known, the parameter development scheme for the iron complex yielded force constants and equilibrium bond lengths for all bonds and angles involving the metal center. Further, the charges on the complex center and its immediate binding atoms were obtained. The vdW parameters for the iron center were adopted from the CHARMM force field for heme. In the following an overview of the parameterization scheme is presented, before selected steps will be elaborated in detail:

1. Build simplified model of the complex (e.g. metal center + functional groups of the ligands).
2. DFT Geometry optimization (B3LYP/6-31+G*).
3. Take vdW parameters for metal center from the existing force field. Possibly missing parameters for the transition metal can be substituted from the Universal Force Field.
4. For ligands, bonded parameters and initial set of atomic partial charges are taken from the existing force field.
5. Compute NPA charges for the complex and for isolated ligands. Assign NPA charge to the metal center. The difference of the total NPA charge of a ligand in its isolated and complexed form is a measure for the charge transfer from the metal center to the ligand. Subtract the transfer charge from the charge of the ligand atom that is bound to the center.
6. Obtain Hessian matrix from frequency analysis.
7. Transform Hessian from cartesian to redundant internal coordinates. The diagonal elements are the force constants.
8. Refine force constants for bonded interactions.

3.2.3 Charges

A new method based on natural population analysis (NPA) [Reed et al., 1985] was developed to determine the charge of the metal center and to estimate the charge transfer from the center to the ligands. The aim was to take the CHARMM charges for the unbound ligands and adjust them in order to account for the effects of binding. First, the NPA charge for iron and the total NPA charge on each ligand was determined. Subtracting the formal charge of the ligand, i.e. –1 for acetate and hydroxide and zero for imidazoles, yielded the net charge transfer from the center to each ligand. These transfer charges were added to the existing CHARMM charges of the ligand atoms that are directly bound to iron.

3.2.4 Bonded Interactions

Based on the optimized geometry the following strategy was pursued to obtain the bonded parameters: The cartesian Hessian was computed for the optimized geometries and subsequently transformed into redundant internal coordinates. The mathematical procedure of the coordinate transformation is given in Appendix C. The following internal coordinates were defined for the transformation: All covalent bonds that occur in the system including the six coordinative bonds between the metal center and the ligands and all bond angles (except linear bends of opposite ligands) and dihedral angles that can be constructed from the bonds. Dihedrals involving iron were omitted, though, except for the ones spanning from Ter663 to hydroxide. For the histidine ligands improper dihedrals centered at the nitrogen atoms forming the complex bonds to iron were added.

Frequencies and force constants derived from DFT based calculations will be systematically overestimated compared to experimental values [Rauhut and Pulay, 1995]. One of the reasons for this disparity is that calculated harmonic frequencies are compared with observed anharmonic fundamentals, because anharmonicity usually lowers the freqencies [Pulay et al., 1983]. Since these errors are mainly of systematic nature, several approaches for scaling the harmonic frequencies, depending on the used basis set, have been described [Rauhut and Pulay, 1995, and references therein]. The suggested scaling factor of 0.963 for B3LYP/6-31G* was used to correct the frequencies in the present work.

3.3 Molecular Dynamics Simulations

All MD-simulations were carried out with the program NAMD [Phillips et al., 2005]. They were performed for an NPT ensemble (constant pressure and temperature) except for equilibrations with fixed atoms in which the NVT ensemble (constant volume and temperature) was employed. When using periodic boundary conditions (PBC), the system is surrounded by translated images of itself, so that the system appears continuous (Fig. 3.7). Atoms leaving the unit cell are wrapped around the system boundaries and are re-inserted at the opposite side of the cell. PBC avoid boundary effects and are a prerequisite for the particle mesh Ewald method, a fast electrostatics calculation algorithm used in the MD simulations. A non-orthogonal PBC cell (a=78 Å, b=72 Å, c=102 Å, and

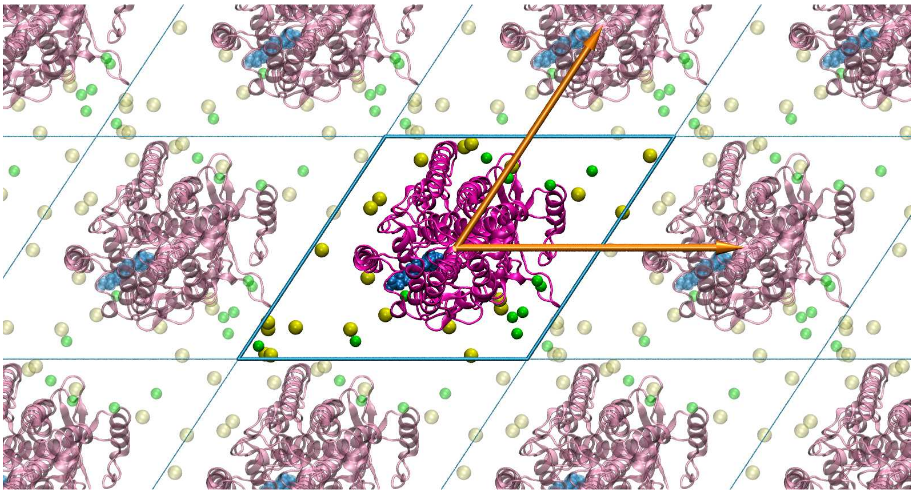


Figure 3.7: Periodic boundary conditions. An infinite, continuous system is generated by virtual copies (shaded) of the PBC cell. The translation vectors are shown in orange.

α=56.6°), enabling optimal packing in space, was used (Fig. 3.5). The size of the unit cell ensured a minimum distance of 15 Å between atoms of adjacent proteins in the lattice.

3.3.1 Equilibration of the enzyme–substrate complex

After completion of basic structural modeling of the enzyme/substrate complex (fitting in the lacking structural LOX components and the fatty acid substrate), the system has to be carefully equilibrated in multiple steps so that the inserted parts can accommodate to the rest. First, all water molecules, the inserted protein parts, the substrate, and the iron complex were energetically minimized and equilibrated for 50 ps, while all other atoms were fixed. Further, the pro-S hydrogen atom at C13 of the arachidonic acid backbone was restrained in bonding distance (2.3 Å) to the OH^--ligand of the ferric iron. Next, the complete protein was released, but for the first 130 ps the modeled helix was constrained to the rest of the protein by an external force of 100 pN to prevent the helix breaking while the contact side chains find their energetically favored positions. This initialization phase was then followed by a 230 ps period, in which only the substrate molecule was restrained to the complex. Finally, the system was equilibrated for 1 ns without any constraints. At the same time, a variant of the system that contains no substrate was generated, following the same protocol. The root mean square deviation of the α-carbon atom positions before and after the equilibration is about 1.0 Å, indicating a stable protein model.

3.3.2 MD simulations of intra-enzyme oxygen movement

After equilibration, intra-enzyme oxygen movement was tested by two independent approaches. To assess oxygen movement within the 12/15-LOX protein, one molecule of dioxygen was placed inside the substrate-binding pocket in close proximity to the site of oxygen insertion during catalysis (C15 of arachidonic acid). Another oxygen molecule was added next to the side chains of H426, V427, and K146, a site homologous to the internal part of the soybean LOX-1 oxygen channel [Knapp and Klinman, 2003]. Oxygen diffusion was monitored over a 2-ns simulation period for the wild-type 12/15-LOX and its L367F mutant.

3.3.3 Implicit ligand sampling

In addition to explicit simulations of oxygen diffusion implicit ligand sampling was employed [Cohen et al., 2006]. With this method, it is possible to compute a three dimensional grid of the Gibbs free-energy cost ΔG for placing a molecule from vacuum into a certain volume element (voxel). In a first step the system was equilibrated for a certain time period without oxygen. Next, oxygen molecules were placed in every trajectory frame on test points of a fine grid throughout the entire system. Averaging over many conformations, the free energy can be computed at every grid position considering the interaction energy of oxygen with the other molecules of the system. The free energy for a volume element centered at position r is defined as

$$\Delta G = -k_{\mathrm{B}}T \ln \sum_{m=1}^{M} \sum_{k=1}^{C} \frac{e^{-\beta \Delta E_{mk}(r)}}{MC} \tag{3.1}$$

where C is the number of positions and orientations tested in each voxel, M is the number of different protein conformations, ΔE_{mk} is the interaction energy for the specified conformation and $\beta = 1/k_{\mathrm{B}}T$. The more samples are taken from different frames, the higher the accuracy of the free-energy calculation.

Free-energy maps for the wild type 12/15-LOX and its L367F mutant were calculated. For this purpose, 8,000 protein conformations were sampled from a 4-ns trajectory, which was obtained under the same general conditions used in the explicit oxygen MD simulations except that no oxygen was present in that system. Each volume element of the free-energy map comprised 1 Å^3, and within each of these voxels oxygen was placed in 40 different rotational orientations at each of 27 different positions on a 3x3x3 subgrid. The 3D free-energy distribution mirrors the probability of oxygen to occupy a given spatial region of the enzyme.

3.3.4 Identification of low barrier pathways in free energy maps

Implicit ligand sampling enables to compute maps of the Gibbs free-energy cost of placing a molecule in a certain volume element of a three dimensional grid. To identify routes for oxygen diffusion from the protein surface to the oxygen binding site at the active center, an algorithm was developed that finds the minimum energy pathways. This is achieved by gradually 'flooding' the energy landscape beginning from the global minimum of the map. Starting with the lowest energy volume element, we identify all neighboring elements with energy lower than the current flood level and add these to the flooded area. This procedure is repeated until no more lower energy elements directly neighboring the already flooded area can be found. Then the flood level is increased. Saddle points are defined by voxels having neighbors from where a new local minimum can be reached by just walking downhill. When a saddle point is found we construct the steepest descent path back into the basin we came from and forward to the new local minimum. Next we start filling the new minimum and iterate the procedure until we finally reach the target, in our case the protein surface. Thus we connect the global minimum and the solvent phase with the lowest barrier pathway. The algorithm is inspired from watershed transformation [Beucher and Lantéjoul, 1979, Meyer and Beucher, 1990], a technique that is being used for segmentation problems in image processing. In its 3D variant it has also many applications in medical imaging [e.g. Grau et al., 2004]. The difference between image segmentation and the path finding problem is that in the latter one is not interested in the detection of watershed lines separating catchment basins but in the saddle points separating the minima.

3.4 Experiments and kinetic modelling

3.4.1 Site directed mutagenesis

The theoretical findings of this work were undermined by accompanying experimental studies. In order to prove the relevance of the proposed main oxygen access channel for enzyme catalysis, it was attempted to obstruct the channel by site-directed mutagenesis. While the interesting sites for mutagenesis were suggested by myself, based on steric considerations of residues along the proposed channel, all mutagenesis experiments were performed by Matthias Walter and Igor Ivanov from Prof. Kühn's group at the Institute of Biochemistry (Charitè). Experimental details are given in appendix C.

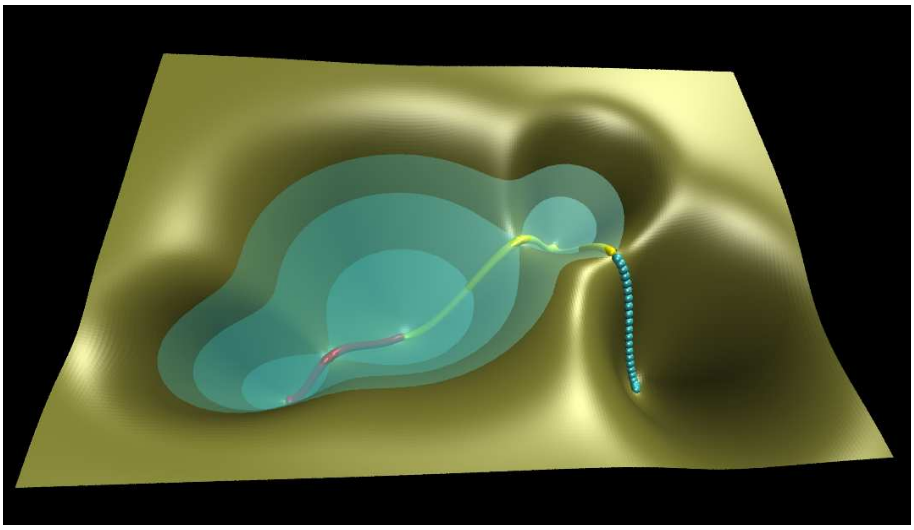


Figure 3.8: Finding paths in free energy maps. The landscape is gradually flooded as indicated by the blue water levels. Whenever a saddle point is detected, the steepest descent paths into the two minima are constructed (red and yellow). Then the new basin is filles (blue dotted line).

3.4.2 Kinetic measurements and quantification of oxygen affinity

For basic kinetic characterization of the mutants the oxygenation of various concentrations of linoleic acid by the different recombinant enzyme species was assayed spectrophotometrically by measuring the time-dependent increase in absorbance at 235 nm. The enzyme was initially activated by incubation with 13S-HpODE (2 μM) in 0.1 M potassium phosphate buffer pH 7.4. The reaction was started by addition of linoleic acid (500 μM) and all spectrophotometric measurements were carried out at room temperature. Different oxygen concentrations were adjusted by mixing hypoxic and hyperoxic assay buffers and the oxygen concentration in the assay system was determined using a Clark-electrode. Fatty acid substrate did never become rate limiting during the time course of oxygenation (pseudo-1st-order conditions). The kinetic measurements were contributed by Igor Ivanov in Prof. Kühn's group.

Quantitative analysis of the kinetic data was performed in close cooperation with Prof. Holzhütter (Institute for Biochemistry, Charité) based on the reaction scheme

$$\mathrm{E} \underset{k_{-1}}{\overset{k_{+1}[\mathrm{S}]}{\rightleftharpoons}} \mathrm{ES} \xrightarrow{k_{\mathrm{S}}} \mathrm{ES}^* \underset{k_{-0}}{\overset{k_{+0}[\mathrm{O}_2]}{\rightleftharpoons}} \mathrm{ES}^*\mathrm{O}_2 \xrightarrow{k_{\mathrm{P}}} \mathrm{P} \tag{3.2}$$

where ES* is the enzyme-fatty acid radical complex and ES*O$_2$ is the complex with channelled oxygen. If the non-stationary activation phase of the reaction is neglected and channelling of oxygen between the active site and the external space is explicitly taken into account, this scheme is identical with the core of a more detailed reaction scheme developed previously Ivanov et al. [2005]. Stationary kinetic treatment results approximately in the rate law

$$v = [\mathrm{E_{tot}}] \frac{k_{\mathrm{cat}} \cdot [\mathrm{O_2}]}{K_{\mathrm{M}}^{\mathrm{O_2}} + [\mathrm{O_2}]} \tag{3.3}$$

where $[\mathrm{E_{tot}}]$ is the total enzyme concentration. The catalytic rate constant is identical with the rate constant for hydrogen abstraction, i.e. $k_{\mathrm{cat}} = k_{\mathrm{S}}$. The ratio between k_{cat} and the apparent Michaelis constant for oxygen is given by

$$\frac{k_{\mathrm{cat}}}{K_{\mathrm{M}}^{\mathrm{O_2}}} = \frac{1}{\frac{1}{K_0 k_{\mathrm{P}}} + \frac{1}{k_{+0}}} \tag{3.4}$$

where k_{P} denotes the rate constant for oxygen insertion into the fatty acid radical. $K_0 = k_{+0}/k_{-0}$ is the equilibrium constant for oxygen and k_{+0} and k_{-0} mean the rate constants for uptake and release of oxygen. Details of the derivation of the above expressions are given in Appendix D.

Numerical values for the parameters $V_{\max} = k_{\mathrm{cat}}[\mathrm{E_{tot}}]$ and $K_{\mathrm{M}}^{O_2}$ were determined by minimizing the sum of squared distances between the rate equation 3.3 and experimentally determined initial rates taken from the slope of photometric progress curves.

Chapter 4

Results

4.1 Force field for the iron complex

4.1.1 Optimized DFT iron-complex geometry

Basis for the parameterization effort is the geometry of the iron complex resulting from DFT optimization. Together with data from X-ray absorption spectroscopy (also: extended x-ray absorption fine structure; EXAFS) [Kuban et al., 1998] it denotes the most detailed model of the equilibrium conformation of the complex currently available for both oxidation states. In Table 4.1 the lengths of all bonds between iron center and ligands are compared. The most pronounced disparity is observed for the bond to His545, which is about 0.8 Å longer in the EXAFS data. The iron–His361 bond lengths are very similar for the ferrous complex, but for ferric iron it is 0.4 Å longer in the EXAFS geometry. All other bond lengths agree within 0.2 Å. Besides through inherent inaccuracies of the DFT calculations, such as the limited basis set, differences in the geometries might be explained by the absence of the protein environment in the quantum mechanical model. Steric constraints imposed by the protein on the ligands and charged groups in the vicinity of the complex may influence the complex geometry. Moreover, the samples containing the ferric form of the complex were subjected to lyophilization (freeze drying), which might also be responsible for structural changes.

4.1.2 Charges

A set of atom centered point charges was obtained from natural population analysis of the DFT optimized molecular geometry. First the NPA-charge (+1.688e) was assigned to the iron atom. The difference of 1.312e to the formal charge of +3 was distributed to the ligands according to the NPA charge transfer that occurs when a free ligand is bound to the complex (Table 4.2). The trans-

	Fe^{3+}			Fe^{2+}		
	k_r	r_QM	r_EXAFS	k_r	r_QM	r_EXAFS
Fe–His361:N_ϵ	45.2	2.26	2.65	44.1	2.24	2.13
Fe–His366:N_ϵ	60.0	2.20	2.25	38.7	2.27	2.13
Fe–His541:N_ϵ	40.5	2.31	2.44	42.8	2.25	2.36
Fe–His545:N_δ	55.7	2.22	3.08	46.2	2.24	2.93
Fe–Ter663:O_1	147.9	1.97	1.92	125.5	2.04	1.89
Fe–OH/H_2O:O	220.1	1.83	2.08	45.5	2.24	2.42

Table 4.1: Geometry of iron-complex. Comparison of iron–ligand bond lengths for B3LYP/6-31+G* (r_QM) with bond lengths from EXAFS (r_EXAFS). The force constants k_r for these bonds are also given.

fer charges are added to the existing CHARMM charges of the ligand atoms involved the complex bonds. The largest transfer of 0.5e occurs from the hydroxide anion, acetate donates about 0.3e, while the depletion at the imidazoles amounts approximately 0.1e.

	formal charge	charge transfer
His361	0	+0.121
His366	0	+0.142
His541	0	+0.097
His545	0	+0.137
Ter663	-1	+0.315
OH	-1	+0.502

Table 4.2: NPA based charge transfer from iron into the ligands.

4.1.3 Bonded interactions

The cartesian Hessian yielded from DFT frequency calculation was transformed to the set of redundant internal coordinates defined by the bonds, angles and dihedrals of the model system. The diagonal elements from the internal Hessian denote the force constants for the associated vibrations, while the equilibrium distances/angles are given by the optimized geometry.

However, if the force constants for bonded interactions are taken directly from the Hessian one introduces an error due to implicit double accounting of electrostatic interactions. DFT based force constants reflect all molecular forces implicated with the according motion, including nonbonded contributions. Considering a bond, for instance, the DFT force constant includes the electrostatic interaction between the moieties at both sides of the bond. However, in the force field there are separate terms for nonbonded interactions. Hence, the nonbonded contribution should be removed from the force constant.

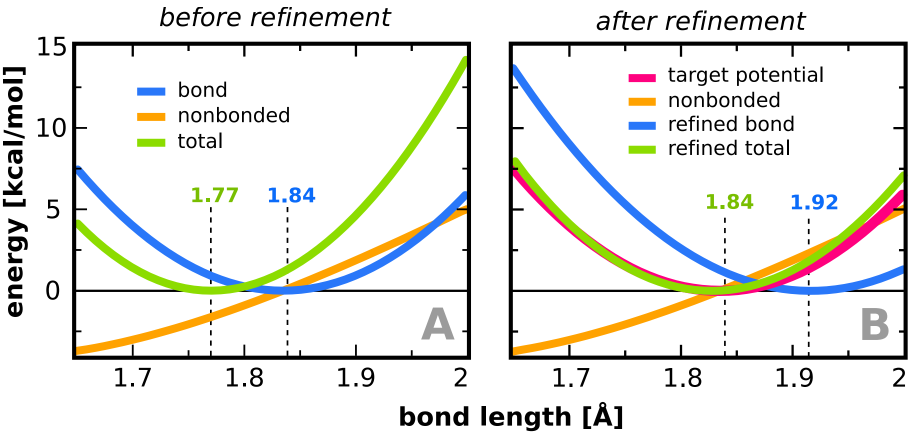


Figure 4.1: Parameter refinement. Potential energy versus bond length for the Fe–OH bond in the iron complex. (A) Harmonic bond stretching term is depicted in blue ($x_0 = 1.84$ Å, $k = 220$ kcal mol^{-1}Å^{-2}). Together with the nonbonded terms (orange) this yields a total potential energy (green) with minimum the at 1.77 Å. The curves for nonbonded and total energy were shifted vertically that the root matches with the bond curve or the minimum is zero, respectively. (B) The bond curve with DFT parameters was defined as the refinement target (red). The refined bond potential in blue ($x_0 = 1.92$ Å, $k = 193$ kcal mol^{-1}Å^{-2}) is shifted, leading to a total energy (green) that agrees better with the target.

Figure 4.1 illustrates the problem for the bond between iron and the hydroxide ligand. In panel A, the blue curve shows the contribution of the bond stretching term to the potential energy of the force field under the assumption that the DFT bond length and force constant is used. The attractive nonbonded contribution (orange) due to the charges on hydroxide and other ligands shifts the resulting total energy (green) towards shorter bond lengths. This means that the equilibrium geometry in a force field based model will exhibit a shorter bond length than the DFT geometry. In panel B the bond force constant is adjusted in a way such that the resulting total energy matches the target potential, i.e. the harmonic approximation of the DFT potential. All bonds between iron center and the ligands were refined in this manner, thus the bond lengths in Table 4.1 and 4.3 differ.

A similar approach was applied for angles, in order to correct for the implicit double accounting of nonbonded interactions. Urey-Bradley terms were added for iron-centered angles, shifting the equilibrium values towards the target value. These angles are interdependent, i.e. changing one of the angles implies a change in other angles, too. For instance, if the angle between two ligands is increased, the angle to their neighboring ligands in the same plane will decrease. In these cases, the target potential for an angle is a combination of the potentials of all simultaneously changed angles. Due to the mutual dependence of the angles, several iterations of refinement were necessary.

The only ligands with a fixed rotational relationship are Ter663 and hydroxide, which is asserted by fairly stiff dihedrals involving both groups. Four manually introduced improper dihedral angles keep the imidazole rings in line with the complex bonds. The force constants were manually set to 20.0 kcal mol^{-1} rad^{-2}, because the DFT values were too weak to prevent the complex from collapsing.

The quality of refined bonds and angles was assumed sufficient when the force constants were within 15% of the target and the equilibrium distance or angle was within 0.05 Å or 2° of the target. The average structure of a 200-ps MD test simulation reproduced the correct DFT geometry with a root mean square deviation of 0.1 Å. Table 4.3 lists the force constants and equilibrium geometries of all bonded interactions that were added to the existing CHARMM force field.

4.2 Analyzing free-energy maps

4.2.1 Dioxygen is concentrated at the catalytic center

Employing implicit ligand sampling [Cohen et al., 2006] we calculated a 3-dimensional distribution map of the Gibbs free energy $\Delta G(\mathrm{O}_2)$ for placing one molecule of dioxygen from vacuum into any 1 Å^3 volume element. In regions where $\Delta G(\mathrm{O}_2)$ is low the probability for finding oxygen, i. e. the oxygen affinity, is high. In water this free energy distribution is uniform (8.2 kJ/mol). In the subsequent text and figures we are using the free energy difference with respect to the solvent instead of vacuum $\Delta G_\mathrm{S}(\mathrm{O}_2)$. In Fig. 4.2 the free energy distribution in the environment of the empty substrate binding pocket is displayed in form of four nested energy isosurfaces. The inner surfaces enclose regions with lower $\Delta G_\mathrm{S}(\mathrm{O}_2)$. Localization of the substrate-binding pocket was visualized by an arachidonic acid molecule albeit the structural model used for these simulations did not contain fatty acid substrate. The free energy maps indicate that the innermost regions of the substrate-binding pocket exhibit a higher probability for finding oxygen than other parts of the enzyme. The global free energy minimum

bond	k_r	r_0		
Fe - His361:N$_\epsilon$	55.8	2.55		
Fe - His366:N$_\epsilon$	63.5	2.45		
Fe - His541:N$_\epsilon$	41.3	2.62		
Fe - His545:N$_\delta$	57.5	2.50		
Fe - Ter663:O$_1$	62.7	1.91		
Fe - OH:O	192.9	1.92		
angle	k_θ	θ_0	k_{UB}	s_0
His361:N$_\epsilon$ - Fe - His366:N$_\epsilon$	28.2	86.5	5.6	3.32
His361:N$_\epsilon$ - Fe - His541:N$_\epsilon$	41.5	86.4	3.2	3.50
His361:N$_\epsilon$ - Fe - His545:N$_\delta$	25.2	90.0	3.2	3.37
His361:N$_\epsilon$ - Fe - OH:O	28.2	88.5	22.6	3.08
His366:N$_\epsilon$ - Fe - His541:N$_\epsilon$	30.4	88.4	0.8	2.87
His366:N$_\epsilon$ - Fe - Ter663:O$_1$	22.1	90.6	5.2	2.74
His366:N$_\epsilon$ - Fe - OH:O	26.5	94.2	0.1	5.15
His541:N$_\epsilon$ - Fe - His545:N$_\delta$	40.7	86.0	2.2	3.46
His541:N$_\epsilon$ - Fe - Ter663:O$_1$	41.7	84.9	3.1	3.51
His545:N$_\delta$ - Fe - Ter663:O$_1$	21.9	92.0	3.9	2.85
His545:N$_\delta$ - Fe - OH:O	33.2	91.2		
Ter663:O$_1$ - Fe - OH:O	61.7	99.9	28.1	2.56
Fe - OH:O - OH:H	32.2	107.6		
Fe - O$_1$ - Ter663:C	121.6	127.7		
dihedral	k_Φ	δ	n	
Fe - O$_1$ - C - Ter663:C$_\alpha$	7.5	180	2	
Fe - O$_1$ - C - Ter663:O$_2$	30.0	180	1	
OH:O - Fe - O$_1$ - Ter663:C	30.0	180	1	
OH:H - OH:O - Fe - Ter663:O$_1$	14.0	180	1	
improper	k_χ	χ_0		
His361:N$_\epsilon$ - Fe - His361:C$_\delta$ - C$_\epsilon$	20.0	0		
His366:N$_\epsilon$ - Fe - His366:C$_\delta$ - C$_\epsilon$	20.0	0		
His541:N$_\epsilon$ - Fe - His541:C$_\delta$ - C$_\epsilon$	20.0	0		
His545:N$_\delta$ - Fe - His545:C$_\gamma$ - C$_\epsilon$	20.0	0		

Table 4.3: Force field parameters. k_r in kcal mol^{-1}Å^{-2}, k_θ, k_Φ, k_χ in kcal mol^{-1} rad^{-2}, bond length r_0 in Å, all angles in degrees.

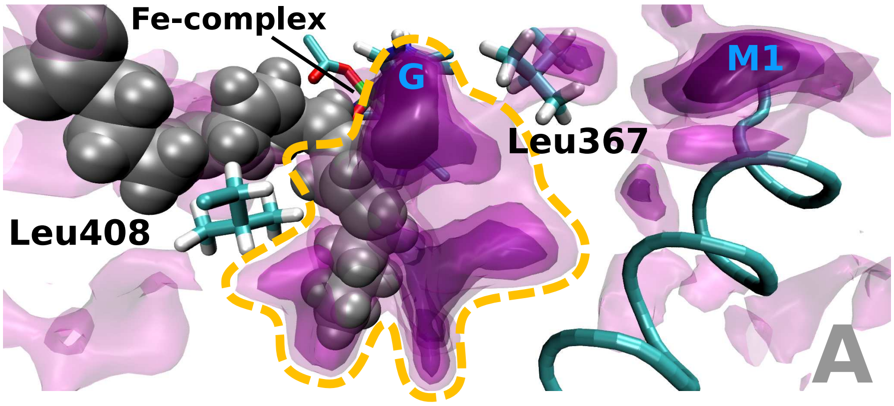


Figure 4.2: Free energy distribution for oxygen inside the substrate-free rabbit 12/15-LOX. (A) Four nested free energy isosurfaces with energy levels of −12.1, −9.5, −7.0 and −3.0 kJ/mol (from dark to light violet) are superimposed. The dark regions represent areas with high probability of finding oxygen (low energy). Although this model represents the substrate free enzyme, arachidonic acid (grey) was projected in to allow better orientation. Parts of the iron-complex can be seen behind the substrate molecule. In front of arachidonic acid the global maximum of the probability of oxygen occupation (G) is located. The central high affinity region close to the active site which is referred to in the text is indicated by a yellow dashed line.

(−19.8 kJ/mol) is located deeply inside the substrate-binding pocket opposite to the non-heme iron. The chance of finding oxygen in the high affinity region around the active site (marked by a yellow dashed line in Fig. 4.2) was 50–100-fold higher than in an equivalent volume of solvent. It has to be noted that these values represent statistical averages sampled from the oxygen distribution pattern of many 12/15-LOX conformations.

4.2.2 Major routes for oxygen movement in substrate-free LOX

Inspection of the highest energy isosurface in Fig. 4.2 at −3.0 kJ/mol provides an impression of possible oxygen pathways. However, to identify the most likely routes for oxygen diffusion from the protein surface to the high affinity area at the active center we developed an algorithm to search for low energy paths in the 3-dimentional energy map. This was accomplished by gradually flooding the energy landscape. Whenever an outlet from the current basin is found a steepest descent path through this saddle point connecting the current with the neighboring basin can be defined. Following this procedure for all newly found

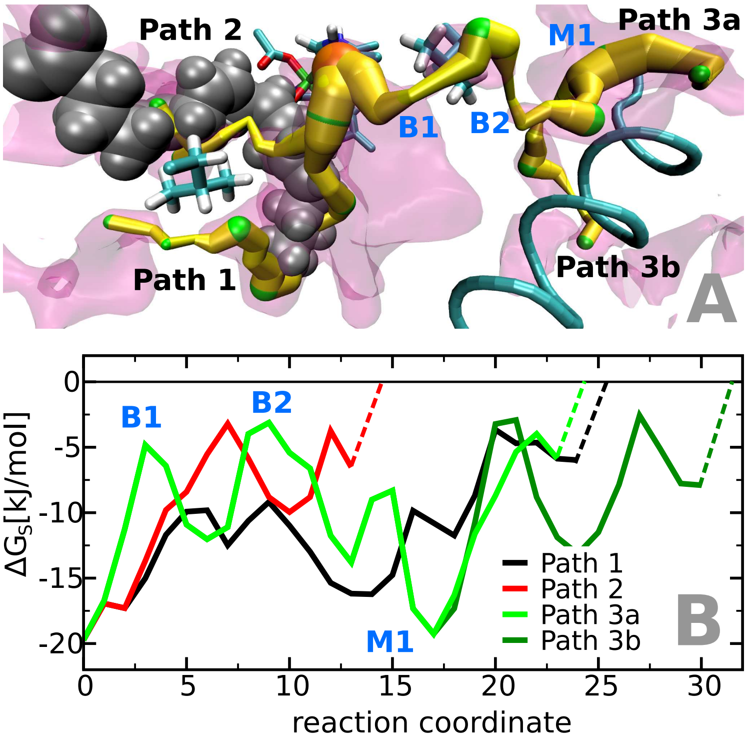


Figure 4.3: Oxygen paths inside the substrate-free rabbit 12/15-LOX. (A) Optimal routes for oxygen movement determined by the "flooding" method. The depicted paths start at the global energy minimum (orange) and end at local minima at the protein surface which are in contact with the solvent. The diameter of the paths is related to the free energy level and thus the oxygen affinity. A thick path symbolizes high oxygen affinity. Local energy minima are marked in green. The inner part of path 1 coincides with the deepest section of the substrate binding pocket. The protein surface is intersected in the surrounding of L408. Path 2 is identical with the substrate entrance while path 3 leads to the opposite side of the protein. Just below the protein surface we identified a deep local energetic minimum (M1) that may act as initial scavenger for oxygen molecules freely diffusing in the solvent. This initial sink can be accessed from either side of a helix formed by residues 426–438. (B) Energy profiles for oxygen movement along the different paths. The dashed lines connecting the profiles with the energy level of oxygen in the solvent represent the movement of oxygen molecules from the solvent into the initial oxygen traps at the protein surface. The reaction coordinate on the abscissa indicates the length of the oxygen paths in arbitrary units. In the surrounding of L367 path 3 has two characteristic energy barriers (B1, B2) suggesting that this amino acid is critical for oxygen movement along this path.

local minima in an iterative manner allows to construct pathways in the free energy landscape that can be considered energetically preferred routes of oxygen diffusion.

Using the substrate-free model of 12/15-LOX we identified three major channels with high probability of oxygen occupancy interconnecting the protein surface with the high affinity area at the active site (Fig. 4.3A). The energy profiles along these paths are shown in Fig. 4.3B. The route of oxygen movement with the lowest barrier (path 1) starts at the bottom of the substrate-binding pocket and reaches the protein surface close to L408 and I414. The inner part of this path corresponds to the oxygen channel postulated for the soybean LOX-1. The second route (path 2) follows the fatty acid-binding pocket. In the absence of the lipid substrate this route is free of major energetic or steric constraints for oxygen movement. The third channel (path 3) connects the opposite side of the protein molecule with the active site. The energy profile along this path is characterized by two barriers related to the side chain of L367 and a deep local energy minimum (M1) between I150, M368, L437 and L508 just below the protein surface. This minimum can be accessed from both sides of the surface helix formed by residues 425–436. In contrast to the substrate-binding pocket path 3 is not continuously open and thus escapes detection when one inspects the crystal structure. It rather resembles a chain of separate, mostly hydrophobic cavities that are transiently interconnected owing to side chain flexibility of lining amino acids.

4.2.3 Arachidonic acid closes the substrate binding pocket for oxygen diffusion but opens a fourth oxygen access channel

When arachidonic acid was added to the system (Fig. 4.4A) the regions of highest oxygen affinity at the active site were not significantly altered and the location of the global energy minimum remained the same. Arachidonic acid does not completely fill the volume of the deeper parts of the binding pocket leaving space for oxygen. This result indicates that dioxygen can be stored at the active site independent of bound substrate. However, towards the substrate entrance where the binding pocket is narrow, fatty acid substrate displaces oxygen from the pocket. Thus, diffusion of oxygen through the substrate-binding pocket is interrupted in the presence of fatty acid substrate. Fig. 4.4B shows that path 1 also disappeared. On the other hand, a fourth oxygen path was opened which reaches the protein surface between W145 and H426 while path 3 is persistent.

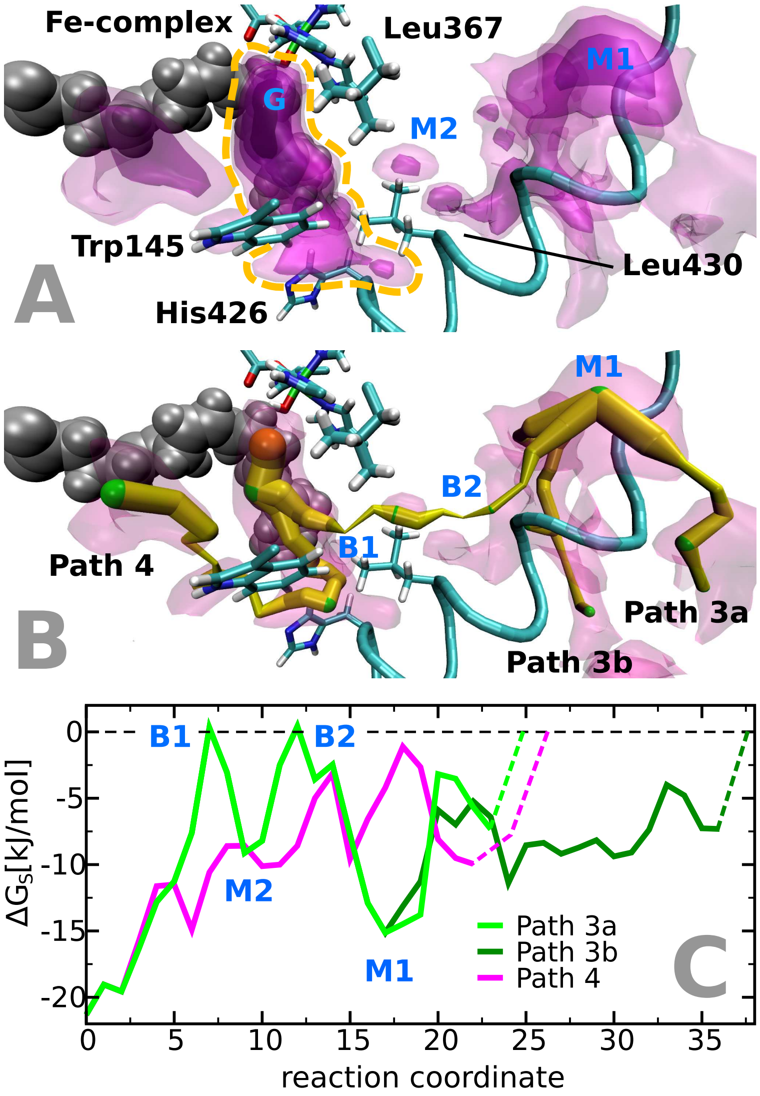


Figure 4.4: Free energy distribution of oxygen inside the 12/15-LOX-arachidonic acid complex. (A) Four nested isosurfaces characterizing the energy levels for oxygen distribution at levels −1.0, −4.9, −10.8 and −16.0 kJ/mol (see legend to Fig. 4.2) are shown. Arachidonic acid is given in grey, the high affinity region at the active site is marked by a yellow dashed line. To visualize the steric relations in a better way this view is rotated right and up when compared with Fig. 4.2. (B) Paths 1 and 2 from the substrate-free map disappeared but a new channel (path 4) reaching the protein surface between W145 and H426 becomes apparent. Panel C: Energy profiles for oxygen movement along paths 3 and 4. In path 3 the two energy barriers are somewhat higher when compared with the substrate free enzyme.

4.2.4 Channel conductivity depends on the energy profile along the path

The energy profiles for oxygen movement along the four major oxygen access channels found in the free energy landscape are shown in Figs. 4.3B and 4.4C. The entrances to all channels are formed by solvent exposed invaginations of the protein surface which exhibit a much higher oxygen affinity when compared to the solvent. The fact that oxygen falls from the solvent energy level into the entrance is symbolized by the dashed lines. Consequently, these hydrophobic invaginations may serve as sinks for initial oxygen enrichment at the protein surface. On their way to the global energy minimum (-19.8 kJ/mol) at the catalytic center oxygen molecules must overcome several energy barriers. Oxygen conductivity along the putative channels depends on both the extent of the energy barriers to be overcome and local increase of oxygen density near the surface invaginations. For instance, path 3 involves high energy barriers that impair oxygen conductivity. On the other hand, the local energy minimum M1 at the protein surface is very low (-19.2 kJ/mol) suggesting high local oxygen concentrations.

Comparing the energy profiles along path 3 of the substrate free and substrate containing 12/15-LOX we found that the internal energy barriers of the enzyme-substrate complex are more pronounced. The molecular changes accounting for this effect have not been studied in detail but may be related to reduced side chain flexibilities in the surrounding of the substrate-binding pocket.

4.3 Implications for catalysis

4.3.1 Directed oxygen access fosters positional and stereospecificity of the reaction

Since the fatty acid radical intermediates might adopt a different conformation in the pocket some care has to be taken relating the oxygen distribution to the specificity of the oxygenase reaction. However, the region where intra-enzyme oxygen concentration is maximal and the iron complex are located roughly at opposite sides of the fatty acid backbone. This spatial arrangement is consistent with the antarafacial character of the LOX reaction, which means hydrogen abstraction and oxygen insertion are taking place from opposite directions of the fatty acid molecule. Furthermore, it can be seen that the oxygen occupation probability within a sphere with 2Å radius around C15 of the arachidonic acid backbone is 7 fold higher than around C11. Thus oxygen insertion at C15 appears to be favored over C11, which is consistent with the positional specificity of the enzyme. However, owing to the high degree of structural flexibility of arachidonic acid our calculations do not exclude oxygen insertion at C11 or at the pro-R side of C15 so that additional mechanisms may contribute to determine the stereochemistry of the oxygenation process [Coffa and Brash, 2004].

4.3.2 High oxygen affinity at active center is important for effective catalysis

A saturation concentration of 280 μM corresponds to one oxygen molecule in $5.9 \cdot 10^6$ Å^3 (a box of 181 Å side length). The oxygen high affinity region at the active site of the substrate loaded 12/15-LOX has a volume of 246 Å^3. If this region had the same oxygen affinity as the solvent only every 20,000th enzyme molecule would be carrying oxygen. Integrating the occupation probability over all volume elements of the high affinity region yields an O_2 occupancy at the active center that is 100-fold higher than the average occupancy of the same volume in the solvent. It follows that in equilibrium about every 200th enzyme molecule is loaded with oxygen. These numbers illustrate that, in order to maintain a rate of 10 s^{-1} with only 1/200 enzymes in action at any instant the product formation could last no longer than $1/(200 \cdot 10)\text{s} = 500\mu\text{s}$.

4.4 L367F exchange corrupts oxygen access

Since in the MD calculations only path 3 appeared to be functional for both substrate-free 12/15-LOX and the enzyme-substrate complex it was attempted to reduce its conductivity by site-directed mutagenesis. For this purpose L367, which lines the channel at a critical position, was mutated to residues carrying more space-filling or charged side chains (L367F, L367W, L367K, L367E). All mutants exhibited a somewhat reduced catalytic activity but other enzymatic characteristics were not altered (see Appendix B).

To assess the effect of the mutations on the oxygen conductivity of path 3 k_{cat}, $K_M^{\mathrm{O_2}}$ was determined and $k_{\mathrm{cat}}/K_M^{\mathrm{O_2}}$ was calculated, the latter revealing specifically how mutants affect oxygen diffusivity (see Table 4.4). For the wild-type enzyme a low $K_M^{\mathrm{O_2}}$ value of 5.2 ± 2.4 μM was obtained indicating a high oxygen affinity in line with previous results [Juranek et al., 1999]. Mutant L367K exhibited a notable drop in k_{cat} and thus was ruled out from further considerations. For the mutants L367F, L367E and L367W the catalytic efficiency was in the order of the wild type but $k_{\mathrm{cat}}/K_M^{\mathrm{O_2}}$ was significantly lowered. The largest changes were observed for L367F showing a 20-fold lower $k_{\mathrm{cat}}/K_M^{\mathrm{O_2}}$ and a 10-fold higher Michaelis constant for oxygen. To interpret these findings with respect to oxygen diffusivity we used the kinetic expressions 3.3 and 3.4 derived for the reaction scheme 3.2. As the calculations yielded very similar Gibbs free energy distributions in the high affinity region of wild type and mutant L367F a significant change of the association constant K_0 for oxygen channeling is unlikely. One can also assume that the rate constant k_P for oxygen insertion into the fatty acid radical stays approximately the same. Under these premises it follows from expression 3.4 that the ratio of the on-rate constants k_{+0}^{wt} and k_{+0}^{mut} for oxygen uptake in the mutant and the wild type enzyme is given by

$$\frac{k_{+0}^{\mathrm{wt}}}{k_{+0}^{\mathrm{mut}}} = 1 + k_{+0}^{\mathrm{wt}} \left(\left(\frac{K_M^{\mathrm{O_2}}}{k_{\mathrm{cat}}} \right)^{\mathrm{mut}} - \left(\frac{K_M^{\mathrm{O_2}}}{k_{\mathrm{cat}}} \right)^{\mathrm{wt}} \right) \tag{4.1}$$

With a typical value of $100\mu\mathrm{M}^{-1}\mathrm{s}^{-1}$ for the diffusion on-rate constant one gets $k_{+0}^{\mathrm{mut}}/k_{+0}^{\mathrm{wt}}$ ratios of 677, 390 and 188 for the mutants L367F, L367E and L367W. Thus, in the mutant enzymes the conductivity of the oxygen channel should be reduced by at least two orders of magnitude.

The charge of the glutamate mutant might account for a reorientation of neighboring sidechains disturbing the architecture of the channel. The effect of the bulky mutants L367F and L367W might to some extent be explained by the reduced channel volume. Mechanistic details accounting for the impact of L367F exchange on oxygen conductivity of channel 3 are revealed when one considers the protein dynamics and the free energy distribution: In Fig. 4.5 the oxygen accessible areas of the substrate-free 12/15-LOX and its L367F mutant are compared in terms of the isosurface at -1.0 kJ/mol. Since the energy level of this isosurface is higher than the largest barrier in the wild type enzyme a continuous channel can be seen (Fig. 4.5A). For the mutant, however, the channel is clearly interrupted and it is not the F367 side chain itself which is blocking the pathway (Fig. 4.5B). Instead, the effect is more subtle. We identified several water molecules in the vicinity of L367 which contribute to opening and closure of the oxygen channel. In the wild-type enzyme these water molecules are not constrained to defined hydration sites but may move around and change their orientations (molecules colored in red, yellow and green in Fig. 4.5). In contrast, for the mutant enzyme (Fig. 4.5B) the water molecules W1 and W2 are immobilized forming a stable hydrogen bonding network that involves E370 and R405. Fig. 4.5B clearly shows that water W1 is blocking the channel. The decreased water mobility is accompanied with the lack of conformational flexibility of the phenylalanine side chain. While L367 is able to rotate more or less freely F367 remained in the same conformation throughout our simulations.

4.5 MD simulation of explicit oxygen diffusion

For further evidence on the paths of intra-enzyme oxygen movement additional MD simulations were performed. They were initialized by placing two oxygen

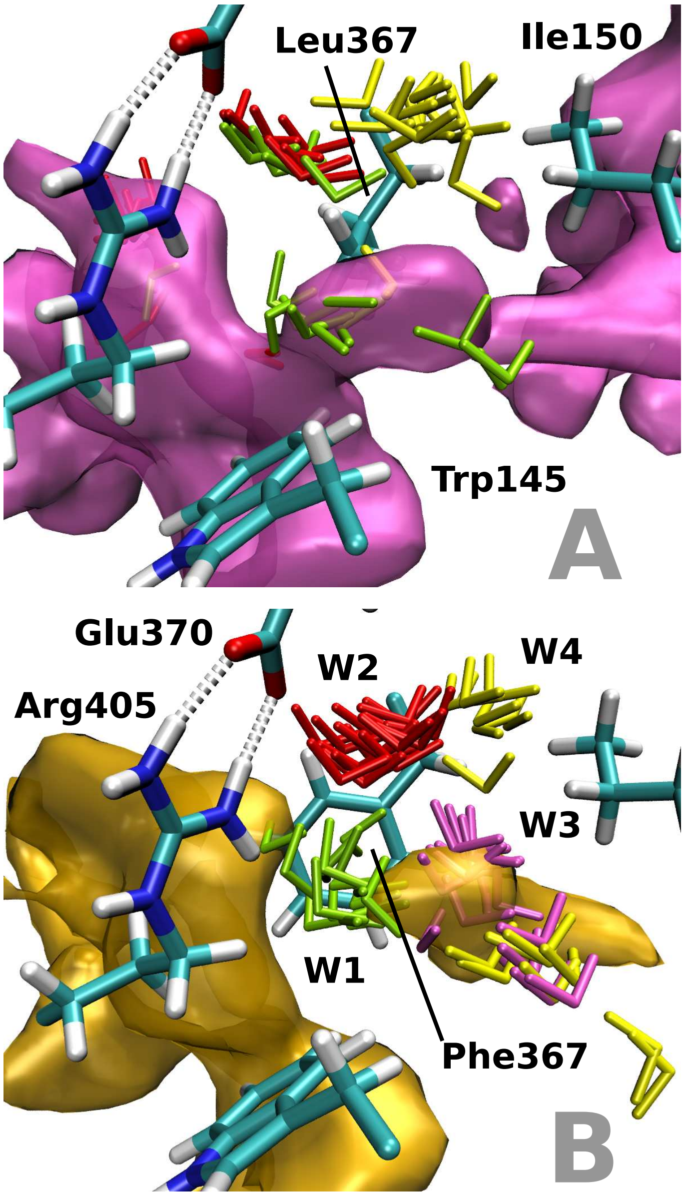


Figure 4.5: Impact of site directed mutagenesis of L367F on intra-enzyme oxygen movement. (A) Free energy isosurface area at −1.0 kJ/mol for the wild type enzyme in the vicinity of L367 (path 3). Three different water molecules (green, red, yellow) in the surrounding of the mutated residue are displayed. Sixteen conformations from a 4 ns trajectory illustrate the mobility of these water molecules. (B) For the L367F mutant the energy isosurface is disrupted at the site of mutation. The water molecules are immobilized and W1 appears to interrupt the flow of oxygen. E370 and R405 participate in the hydrogen-bonding network arresting the water molecules.

Table 4.4: Enzyme characteristics of different mutants

	k_{cat} $[\mathrm{s}^{-1}]$	$k_{cat}/K_{\mathrm{M}}^{\mathrm{S}\star}$ $[\mathrm{s}^{-1}\mu\mathrm{M}^{-1}]$	$K_{\mathrm{M}}^{\mathrm{O}_2}$ $[\mu\mathrm{M}]$	$k_{cat}/K_{\mathrm{M}}^{\mathrm{O}_2}$ $[\mathrm{s}^{-1}\mu\mathrm{M}^{-1}]$
wild type	13.7±0.4	0.74±0.08	5.2±2.4	2.63±1.22
L367K	0.3±0.01	0.08±0.01	9.0±1.8	0.03±0.01
L367E	2.2±0.2	0.26±0.06	9.0±2.1	0.24±0.06
L367F	5.6±0.4	0.73±0.07	40.1±3.8	0.14±0.02
L367W	4.4±0.2	0.48±0.05	7.0±3.2	0.63±0.29

* S = linoleic acid

molecules into the high oxygen affinity area of the wild-type 12/15-LOX with and without bound substrate and the arachidonic acid containing L367F mutant. Since the low barrier paths 1 and 2 are blocked by the fatty acid oxygen molecules do not leave the high affinity area during the 2-ns test simulation along those ways. However, one of the introduced oxygen molecules escaped via path 3 from the wild type enzyme into the surrounding solvent (Fig. 4.6). For the mutant enzyme outward movement of oxygen via path 4 was observed. All other O_2 molecules remained inside the protein. The oxygen trajectories are in accordance with the computed free energy maps, i.e. the oxygen molecules diffuse mainly in regions corresponding to low energy in the maps. Thus, the oxygen channels indeed constitute traps for oxygen. As oxygen hits the barriers many times during the simulations before it eventually surmounts them, the transient and dynamic character of the channels opening only from time to time is illustrated. Of course the number of MD simulations is too small for proper statistics, but the fact that oxygen was able to leave the channel after about one nanosecond is at least interesting. It corresponds to the assumption that the equilibrium of oxygen between solvent and active site should be fast compared to the turnover time. Otherwise oxygen would be rate limiting for the enzyme reaction not only at very low oxygen concentrations. This is also consistent with reports that oxygen reversibly enters the soy LOX-1 enzyme in the presence of fatty acid substrate [Knapp et al., 2001].

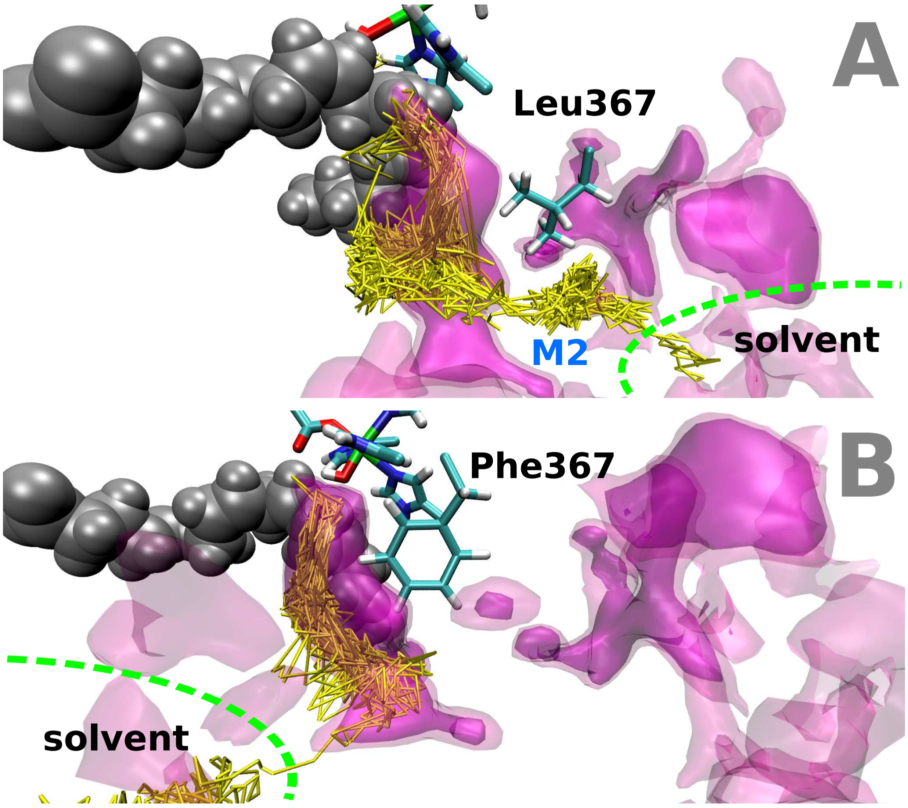


Figure 4.6: Explicit oxygen diffusion trajectories. (*A*) Substrate-free wild-type. Oxygen passes barrier B1 on path 3, stays some time in local minimum M2, explores a cavity that is in contact with the solvent, but finally returns into channel 3. (*B*) L367F mutant. Diffusion inside the region of high oxygen affinity at the catalytic center is shown. Eventually oxygen leaves the protein through path 4.

Chapter 5

Discussion

5.1 Force Field Parameters for the Iron-Complex

There exist several approaches to force fields for transition metal complexes, for a review see [Norrby and Brandt, 2001, Banci, 2003]. The simplest way of treating metal coordination is to use solely nonbonded interactions. This has the advantage of allowing dynamic ligand exchange, but ignores ligand orientation. This limitation does not occur with the 'covalent' model, that defines the bonds between the center and the ligands explicitly, thus fixing the orientation and coordination number which is adequate for reasonably well defined and rigid systems like octahedral complexes. Some of the existing force fields for transition metal complexes [e.g. Vedani and Huhta, 1990, David et al., 2002] use a functional form, that differs from common biomolecular force fields. Hence, they are difficult to use in widespread molecular dynamics packages.

Hoops et al. [1991] have proposed a methodology to derive parameters for metalloproteins that works with the AMBER force field and successfully demonstrated it for a tetrahedral zinc complex with simplified models for the ligands (imidazole rings instead of histidines). Similar approaches were adopted by Suarez and Merz Jr. [2001] for another zinc-complex and Comba and Remenyi [2002] for a 4-coordinated copper complex. These examples have in common, that their charges are based on electrostatic potential (ESP) fits [Singh and Kollman, 1984, Besler et al., 1990] or restricted ESP (RESP) [Bayly et al., 1993] and that the metal centers have only four ligands. At higher coordination numbers, such as in octahedral iron complexes, the geometry has several buried atoms that are statistically underrepresented in the fit [Francl et al., 1996, Stouch and Williams, 1993] and thus cause difficulties with unambiguous charge assignment. Problems regarding the treatment of charges in larger models, i.e. where the central atom is buried deeper, were reported in [Hoops et al., 1991]. However, the use of

RESP may alleviate the buried charge problem, especially for lower coordination numbers. Furthermore, the ESP/RESP fit is conformation dependent, an effect that is pronounced in a dense coordination sphere with close interaction between the ligands. Since the ligands are often able to rotate, a large number of different conformations should be taken into account. Using the CHARMM protocol to assign charges to the iron complex would cause similar problems. The accessibility of hydration sites and the resulting interaction energies would be highly dependent on the rotation of the ligands. In this work a simple solution for the problem of buried charges in the complex using natural population analysis (NPA) is suggested. The advantage of NPA charges is that they are much less conformation dependent. The basic assumption is that the nonbonded interaction properties of a molecule are changed after becoming part of the complex mostly due to the charge transfer from the center to the ligands. The advantage of NPA for estimating the charge transfer is that the charges are much less conformation dependent.

For coordination complexes in proteins the bonded parameters of the isolated ligands are usually known, whereas the interactions with the metal center have to be added. Depending on the changes that are imposed on the ligand itself by the binding to the complex, parameters can be taken from the existing force field or have to be derived anew. In other force fields for metalloproteins [Hoops et al., 1991, Comba and Remenyi, 2002, Suarez and Merz Jr., 2001] it was assumed that these effects on the equilibrium geometry and vibrational frequencies of the ligand are subtle and were omitted. We adopted this approach for the imidazoles in the LOX complex and used the bond, angle and dihedral parameters from CHARMM. For the stronger bound ligands, i.e. hydroxide and acetate the bond, angle and dihedral parameters were derived. The parameters that were introduced to account for the metal–ligand interaction include the bonds between the center and the ligands and the angles that define the ligand positions with respect to each other. The dihedrals of the torsion of ligands around their coordination bond are omitted for the following reason: The torsional potential is mostly due to nonbonded interactions between neighbouring ligands. These interactions are highly interdependent of the orientation of each ligand, hence a simple potential energy scan by rotating around one bond does not make sense. Instead, it was assumed that the rotational interaction is best modeled by vdW and electrostatic contributions.

Main goal of this part of the work was to develop a specific force field parameter set of the iron-complex in 12/15-LOX. However, the technique can be useful for other transition metal complexes in proteins, too, especially when the coordination number is high and the ligands are larger. The parameters were designed as an extension of the CHARMM force field, but due to the similarities of the force fields, the method might be applied as well to GROMACS and OPLS.

The parameters represent a balance between accuracy on the one side and effort and complexity on the other. A more rigorous treatment of slow vibrations might be possible, but since the iron complex is not in the focus of the simulations, the most important property to reproduce is the equilibrium geometry.

5.2 Oxygen access channels in lipoxygenases

For soybean LOX-1 the presence of a putative oxygen access channel has been suggested, but for mammalian LOX isoforms similar data are currently not available. Based on extensive MD simulations and implicit ligand sampling calculations I could predict 4 distinct oxygen access channels joining different sites of the protein surface with a region of high oxygen affinity around the catalytic center. Reliability of the theoretical predictions is supported by the following: i) An Oxygen access channel has been reported before for the soybean LOX-1 [Knapp et al., 2001, Knapp and Klinman, 2003] and we compared the localization of these cavities with those of the rabbit enzyme. The inner oxygen cavity of the soybean enzyme channel co-localizes with the lower part of the substrate-binding cage and with parts of the high oxygen affinity region identified for the rabbit enzyme. Comparison of the more distal regions of the oxygen channels is problematic since these regions do not exhibit a high degree of structural similarity. ii) The mutagenesis experiments on the critical Leu367 residue in channel 3 led to a mutant with significantly reduced oxygen affinity and the molecular reasons for the impaired oxygen permeability have been identified with the help of molecular dynamics simulations (see Fig. 4.5). iii) According to the structural model, the localization of the high affinity region is consistent with the antarafacial character of the LOX reaction. However, the positional specificity of oxygen insertion cannot be explained alone by the distribution of oxygen occupancy probability at C11 and C15 of the arachidonic acid backbone. This is mainly due to the high degree of motional flexibility of the substrate. Thus, alternative mechanisms appear to contribute to the stereocontrol of the LOX reaction [Coffa and Brash, 2004]. Similar conclusions have also been drawn for cyclooxygenase [Furse et al., 2006].

A central finding of our computations is the dynamic character of the predicted oxygen channels some of which transiently open and close due to fluctuations of the protein conformation and thus are not detectable when inspecting the rigid crystallographic protein structure. Furthermore, our calculations show that two of the oxygen channels (path 1 and 2) detected in the substrate free enzyme are blocked upon fatty acid binding. One of them coincides with the substrate binding pocket (path 2). Since hydrogen abstraction is the rate limiting step [Egmond et al., 1973] one can assume an equilibrium of oxygen exchange that is fast compared to the sojourn time of fatty acid in the pocket. Therefore it is unlikely that these two paths play a role in catalysis. In fact, mutation of a critical position along path 1 (L408F and L408W) did not impair oxygen affinity of the enzyme (data not shown). The equilibrium of oxygen exchange in the wild type being much faster than the turnover rate has further implications: An increase of $K_{\mathrm{M}}(\mathrm{O}_2)$ for the mutant indicates that oxygen access is impeded to an extent that oxygen becomes rate limiting.

5.2.1 Detailed comparison of oxygen access channels in soybean LOX-1 and rabbit 15-LOX

An oxygen access channel has been postulated for the structure of soybean LOX-1 (sLOX-1) with linoleic acid modeled-in as substrate [Knapp et al., 2001, Knapp and Klinman, 2003]. The inner part of the channel is formed of a side cavity of the binding pocket surrounded by residues Q495 L546 I553. This cavity coincides with the deepest part of the binding pocket in rabbit LOX (15-rLOX) which may be due to the longer substrate used in our model (residues given for sLOX-1 correspond to E357 L408 F415 in 15-rLOX). At the same time this region overlaps with the oxygen high affinity region from our calculations. After a constriction near V564 and I553 (V427 and F415) the channel opens into another cavity and finally reaches the protein surface close to R203 and P204. Part of the second cavity and the surface intersection lie in nonconserved regions thus comparison in this region is difficult. Figure 5.1 shows the central high affinity region and the oxygen routes of unliganded 15-rLOX together with the residues that define the region corresponding to the channel in sLOX-1.

5.2.2 Occurrence of oxygen channels in other enzymes

Presence of distinct molecular pathways for oxygen transport in proteins has been demonstrated for other enzymes. In water oxygen can freely diffuse and early fluorescence quenching studies suggested that this may also be the case in proteins [Calhoun et al., 1983]. In contrast, more recent studies on oxygen metabolizing enzymes provide examples for targeted oxygen movement. i) For copper amine oxidase reducing O_2 to H_2O_2 when oxidizing primary amines a special channel for oxygen and hydrogen peroxide that is different from the substrate path has been proposed [Li et al., 1998]. ii) Cholesterol oxidase is a bacterial enzyme that oxidizes 3β-hydroxysteroids. In the crystal structure, which is highly resolved (0.95 Å), multiple conformations exist for many of the side chains [Lario et al., 2003, Coulombe et al., 2001]. Based on steric considerations, two conformations for the entire protein could be constructed. In one conformation no channel can

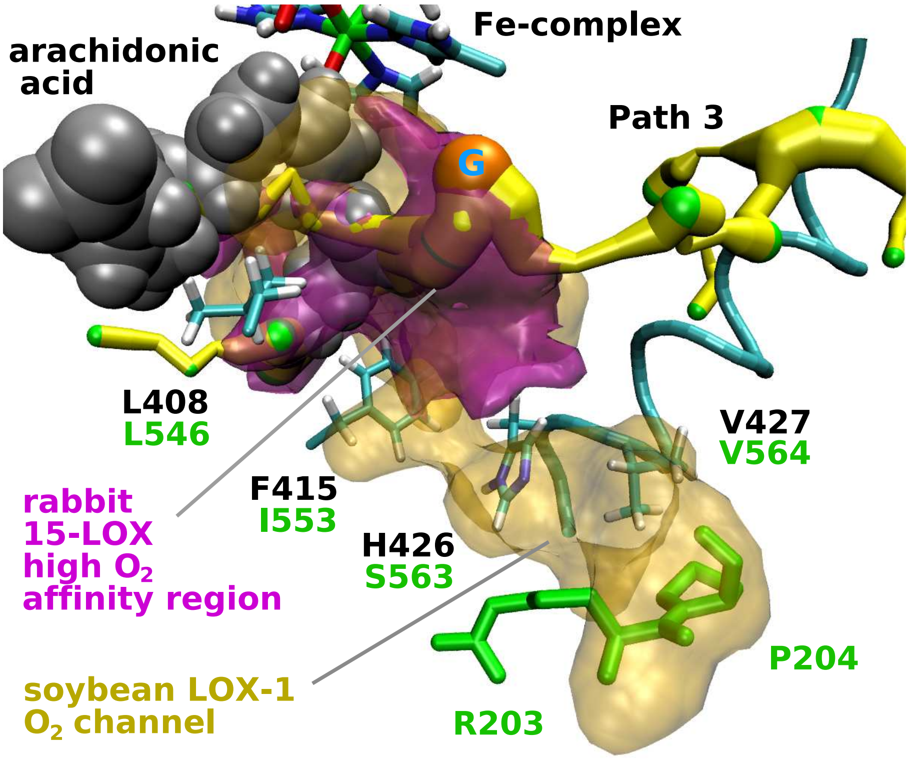


Figure 5.1: Comparison of oxygen channels in soybean and rabbit LOX. Superposition of rabbit 15-LOX oxygen high affinity region (magenta) with putative soybean LOX-1 oxygen channel (ochre). Note that representation of the high affinity region is a free energy isosurface while the sLOX-1 channel visualization is based on the solvent accessible surface of the crystal structure. Residue numbers in green refer to sLOX-1 while the black numbers denote the according residues in the rabbit enzyme. Both models do not contain substrate but arachidonic acid was projected in to visualize the binding pocket. The sLOX-1 channel does not correspond to the 15-rLOX access routes for O_2 but the inner part of the channel colocalizes with the high affinity region.

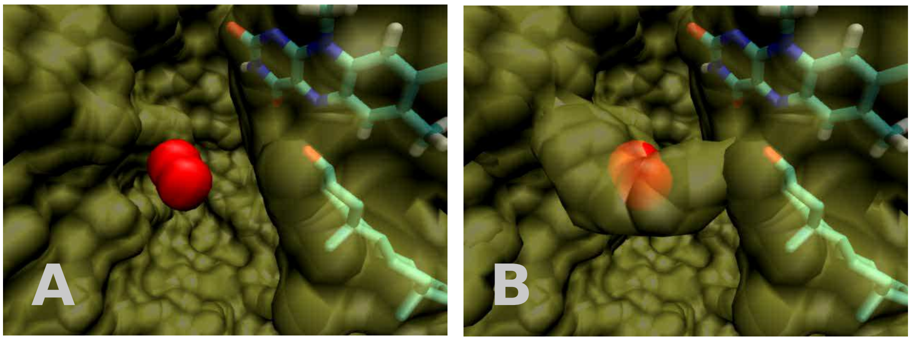


Figure 5.2: Oxygen channel in cholesterol oxidase. Two conformations from the crystal structure are displayed. In the vicinity of the active center an oxygen molecule (red) can be seen. (A) No connection exists between catalytic site (to the right, with cofactor flavin) and the protein surface to the left. (B) In this conformation the sidechains are packed in a different way. A channel opens up.

be seen while in the other there is a hydrophobic tunnel from the protein surface to the active center which is distinct from the substrate binding cleft (Fig. 5.2). This leads to the conclusion that the oxygen channel transiently opens and closes due to side chain dynamics. iii) Cytochrom c oxidase, which catalyzes the final step of the respiratory chain, requires oxygen as electron acceptor. For electron transfer molecular dioxygen must diffuse to the catalytic center located deeply inside the protein. A preformed channel that is supposed to serve as oxygen access path was detected in the enzymes prepared from different species [Riistama et al., 1996, Koutsoupakis et al., 2003] and was confirmed by MD-simulations of oxygen diffusion inside the protein [Hofacker and Schulten, 1998]. iv) For cyclooxygenases solvent-accessible surface calculations and MD-studies of oxygen diffusion inside the binding pocket revealed four access channels to the active site. Interestingly, the channel dynamics suggest that only the route also serving as substrate entrance efficiently functions as oxygen access pathway. This channel provides direct access to the stereochemically correct site of insertion while other possible sites are sterically shielded [Furse et al., 2006]. These literature data and the MD simulations presented in this thesis indicate that oxygen access channels are not necessarily empty or solvent-filled tubes inside the protein. They can be formed of neighboring cavities, which are temporarily interconnected owing to side chain flexibility. Because of their dynamic character such channels might escape detection if static structural models are used.

5.3 Outlook

The free-energy distribution for oxygen tells about the average oxygen occupation probability at all locations in the protein. However, it contains no information about the dynamics of oxygen diffusion. Is there a way to learn about the dynamics in a more efficient way than studying a large number of MD simulations of explicit oxygen diffusion? The free-energy distribution denotes the average potential oxygen is subjected to. Monte-Carlo simulations [Metropolis et al., 1953] of particle diffusion in that potential can give information about relative diffusion times along different routes. Furthermore, they could provide a more detailed picture of the conductivity of different oxygen channels, which is not determined solely by the barrier heights but also by the catchment area of the channel.

In the present work the results of the MD simulations and free-energy calculations in connection with the kinetic and mutagenesis data provide a detailed mechanistic picture and direct experimental evidence for the existence of functional oxygen access channels in the rabbit 12/15-LOX. They prompt the conclusion that similar structures might exist for many other oxygen consuming enzymes. The implicit ligand sampling method can also be applied to other small hydrophobic molecules [Cohen et al., 2006]. In consequence, similar studies assessing the energetics and dynamics of gas diffusion in enzymes utilizing NO, CO or CO_2 appear promising.

Bibliography

B. J. Alder and T. E. Wainwright. Phase transition for hard sphere system. *J. Chem. Phys.*, 27:1208–1209, 1957.

B. J. Alder and T. E. Wainwright. Studies in molecular dynamics i: General method. *J. Chem. Phys.*, 31:459–466, 1959.

L. Banci. Molecular dynamics simulations of metalloproteins. *Curr. Op. Cell Biol.*, 7:143–149, 2003.

C. I. Bayly, P. Cieplak, W. D. Cornell, and P. A. Kollman. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The resp model. *J. Phys. Chem.*, 97:10269–10280, 1993.

A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A*, 38(6):3098–3100, 1988.

H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans. Interaction models for water in relation to protein hydration. In B. Pullman, editor, *Intermolecular Forces*, pages 331–342. D. Reidel, Dordrecht, Holland, 1981.

B. H. Besler, K. M. Merz, and P. A. Kollman. Atomic charges derived from semiempirical methods. *J. Comp. Chem.*, 11:431–439, 1990.

S. Beucher and C. Lantéjoul. Use of watersheds in contour detection. In *International Workshop on Image Processing, Real-Time Edge and Motion Detection/Estimation*, volume 132, pages 2.1–2.12. CCETT/IRISA, 1979.

H. Bisgaard. Pathophysiology of cysteinyl leukotrienes and effects of leukotriene receptor antagonists in asthma. *Allergy*, 56:7–11, 2001.

J. C. Boyington, B. C. Gaffney, and L. M. Amzel. The three dimensional structure of an arachidonic acid 15-lipoxygenase. *Science*, 260:1482–1486, 1993.

A. R. Brash. Lipoxygenases: occurrence, functions, catalysis, and acquisition of substrate. *J. Biol. Chem.*, 274:23679–23482, 1999.

D. B. Calhoun, J. M. Vanderkooi, G. W. Woodrow III, and S. W. Englander. Penetration of dioxygen into proteins studied by quenching of phosphorescence and flourescence. *Biochemistry*, 22:1526–1532, 1983.

R. T. Carroll, J. Muller, J. Grimm, W. R. Dunham, R. H. Sands, and M. O. Funk Jr. Rapid purification of rabbit reticulocyte lipoxygenase for electron paramagnetic spectroscopy characterization of the non-heme iron. *Lipids*, 28:241–244, 1993.

J. Choi, J. K. Chon, S. Kim, and W. Shin. Conformational flexibility in mammalian 15S-lipoxygenase: Reinterpretation of the crystallographic data. *Proteins: Structure, Function, and Bioinformatics*, in press, 2007.

G. Coffa and A. R. Brash. A single active site residue directs oxygenation stereospecificity in lipoxygenases: Stereocontrol is linked to the position of oxygenation. *Proc. Natl. Acad. Sci. USA*, 101:15579–15584, 2004.

J. Cohen, A. Arkhipov, R. Braun, and K. Schulten. Imaging the migration pathways for O2, CO, NO, and Xe inside myoglobin. *Biophys. J.*, 91:1844–1857, 2006.

P. Comba and R. Remenyi. A new molecular mechanics force field for the oxidized form of blue copper proteins. *J. Comp. Chem.*, 23:697–705, 2002.

W. D. Cornell, P. Cieplak, C. I. Bayly, and P. A. Kollman. Application of resp charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation. *J. Am. Chem. Soc.*, 115:9620–9631, 1993.

W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, Jr. K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.*, 117:5179–5197, 1995.

R. Coulombe, K. Q. Yue, S. Ghisla, and A. Vrielink. Oxygen access to the active site of cholesterol oxidase through a narrow channel is gated by an Arg-Glu pair. *J. Biol. Chem.*, 276:30435–30441, 2001.

L. David, P. Amara, M. J. Field, and F. Major. Parametrization of a force field for metals complexed to biomacromolecules: applications to fe(ii), cu(ii) and pb(ii). *J. Comp.-Aided Mol. Design*, 16:635–651, 2002.

M. R. Egmond, G. A. Veldinkand, J. F. Vliegenthart, and J. Boldingh. C-11 h-abstraction from linoleic acid, the rate-limiting step in lipoxygenase catalysis. *Biochem. Biophys. Res. Commun.*, 54:1178–1184, 1973.

R. Elber and M. Karplus. Enhanced sampling in molecular dynamics: Use of the time-dependent Hartree approximation for a simulation of carbon monoxide diffusion through myoglobin. *J. Am. Chem. Soc.*, 112(25):9161–9175, 1990.

O. Ermer. Calculation of molecular properties using force fields. applications in organic chemistry. *Structure and Bonding*, 27:161–211, 1976.

I. Feussner and K. Wasternack. The lipoxygenase pathway. *Annu. Rev. Plant Biol.*, 53:275–297, 2002.

M. M. Francl, C. Carey, L. E. Chirlian, and D. M. Gange. Charges fit to electrostatic potentials ii. can atomic charges be unambiguously fit to electrostatic potentials? *J. Comp. Chem.*, 17:367–383, 1996.

M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople. Gaussian 03 (revision b.05), 2003.

C. D. Funk. Prostaglandins and leukotrienes: Advances in eicosanoid biology. *Science*, 294:1871–1875, 2001.

K. E. Furse, D. A. Pratt, C. Schneider, A. R. Brash, N. A. Porter, and T. P. Lybrand. Molecular dynamics simulations of arachidonic acid-derived pentadienyl radical intermediate complexes with cox-1 and cox-2: Insights into oxygenation regio- and stereoselectivity. *Biochemistry*, 45:3206–3218, 2006.

Q-F. Gan, M. F. Browner, D. L. Sloane, and E. Sigal. Defining the arachidonic acid binding site of human 15-lipoxygenase. *J. Biol. Chem.*, 271:25412–25418, 1996.

J. Garnier, J.-F. Gibrat, and B. Robson. Gor secondary structure prediction method version iv. *Methods in Enzymology*, 266:540–553, 1996.

C. Geourjon and G. Deléage. Sopma: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Comput. Appl. Biosci.*, 11:681–684, 1995.

S. A. Gillmor, A. Villasenor, R. Fletterick, E. Sigal, and M. F. Browner. The structure of mammalian 15-lipoxygenase reveals similarity to the lipases and the determinants of substrate specificity. *Nat. Struct. Biol.*, 4:1003–1009, 1997.

V. Grau, A. U. Mewes, M. Alcañiz, R. Kikinis, and S. K. Warfield. Improved watershed transform for medical image segmentation using prior information. *IEEE Trans. Med. Imaging*, 23:447–458, 2004.

A. T. Hagler. Theoretical simulation of conformation, energetics, and dynamics of peptides. In J. Meienhofer, editor, *Conformation in Biology & Drug Design, The Peptides*, volume 7, pages 213–299. Academic Press, New York, 1985.

T. A. Halgren. Merck molecular force field. i. basis, form, scope, parameterization, and performance of mmff94. *J. Comp. Chem.*, 17:490–519, 1996.

J. Hermans, H. J. C Berendsen, W. F. van Gunsteren, and J. P. M. Postma. A consistent empirical potential for water-protein interactions. *Biopolymers*, 23: 1513–1518, 1984.

I. Hofacker and K. Schulten. Oxygen and proton pathways in cytochrome *c* oxidase. *Proteins: Struct., Func., Gen.*, 30(1):100–107, 1998.

P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136: B864–B871, 1964.

S. C. Hoops, K. W. Anderson, and K. M. Merz Jr. Force field design for metalloproteins. *J. Am. Chem. Soc.*, 113:8262–8270, 1991.

I. Ivanov, S. Romanov, C. Ozdoba, H. Holzhütter, G. Myagkova, and H. Kuhn. Enantioselective substrate specificity of 15-lipoxygenase 1. *Biochemistry*, 43: 15720–15728, 2004.

I. Ivanov, J. Saam, H. Kuhn, and H. Holzhütter. Dual role of oxygen during lipoxygenase reactions. *FEBS J.*, 272:2523–2535, 2005.

W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, 79:926–935, 1983.

W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives. Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.*, 118:11225–11236, 1996.

I. Juranek, H. Suzuki, and S. Yamamoto. ffinities of various mammalian arachidonate lipoxygenases and cyclooxygenases for molecular oxygen as substrate. *Biochim. Biophys. Acta*, 1436:509–518, 1999.

L. Kleinman. Exchange density-functional gradient expansion. *Phys. Rev. B*, 30: 2223–2225, 1984.

M. J. Knapp and J. P. Klinman. Environmentally coupled hydrogen tunneling – linking catalysis to dynamics. *Eur. J. Biochem.*, 269:3113–3121, 2002.

M. J. Knapp and J. P. Klinman. Kinetic studies of oxygen reactivity in soybean lipoxygenase-1. *Biochemistry*, 42:11466–11475, 2003.

M. J. Knapp, F. P. Seebeck, and J. P. Klinman. Steric control of oxygenation regiochemistry in soybean lipoxygenase-1. *J. Am. Chem. Soc.*, 123:2931–2932, 2001.

W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133–A1138, 1965.

K. Koutsoupakis, S. Stavrakis, T. Soulimane, and C. Varotsis. Oxygen-linked equilibrium cub-co species in cytochrome ba3 oxidase from thermus thermophilus. implications for an oxygen channel at the cub site. *J. Biol. Chem.*, 278:14893–14896, 2003.

R. J. Kuban, R. Wiesner, J. Rathmann, G. Veldink, H. Nolting, V. A. Solé, and H. Kuhn. The iron ligand sphere geometry of mammalian 15-lipoxygenases. *Biochem. J.*, 332:237–242, 1998.

H. Kuhn and V. O'Donnel. Inflammation and immune regulation by 12/15-lipoxygenases. *Prog. Lipid Res.*, 45:334–356, 2006.

H. Kuhn, J. Belkner, and R. Wiesner. Subcellular distribution of lipoxygenase products in rabbit reticulocyte membranes. *Eur. J. Biochem.*, 191:221–227, 1990.

H. Kuhn, I. Römisch, and J. Belkner. The role of lipoxygenase isoforms in atherogenesis. *Mol. Nutr. Food Res.*, 49:1014–1029, 2005a.

H. Kuhn, J. Saam, S. Eibach, H. G. Holzhütter, I. Ivanov, and M. Walther. Structural biology of mammalian lipoxygenases: Enzymatic consequences of targeted alterations of the protein structure. *Biochem. Biophys. Res. Commun.*, 338:93–101, 2005b.

J. R. Lakowicz and G. Weber. Quenching of fluorescence by oxygen. a probe for structural fluctuations in macromolecules. *Biochemistry*, 12:4161–4170, 1973a.

J. R. Lakowicz and G. Weber. Quenching of fluorescence by oxygen. detection of structural fluctuations in proteins on the nanosecond time scale. *J. Mol. Biol.*, 12:4171–4179, 1973b.

P. I. Lario, N. Sampson, and A. Vrielink. Sub-atomic resolution crystal structure of cholesterol oxidase: What atomic resolution crystallography reveals about enzyme mechanism and the role of FAD cofactor in redox activity. *J. Mol. Biol.*, 326:1635–1650, 2003.

N. Lehnert and E. I. Solomon. Density-functional investigation on the mechanism of H-atom abstraction by lipoxygenase. *J. Biol. Inorg. Chem.*, 8:294–305, 2003.

R. Li, J. P. Klinman, and F. S. Matthews. Copper amine oxidase from hansenula polymorpha: the crystal structure determined at 2.4 Å resolution reveals the active conformation. *Structure*, 6:293–307, 1998.

R. L. Maas and A. R. Brash. Evidence for a lipoxygenase mechanism in the biosynthesis of epoxide and dihydroxy leukotrienes from 15(s)-hydroperoxy-icosatetraenoic acid by human platelets and porcine leukocytes. *Proc. Natl. Acad. Sci. USA*, 80:2884–2888, 1983.

A. D. Mackerell. Empirical force fields for biological macromolecules: Overview and issues. *J. Comp. Chem.*, 25:1584–1604, 2004.

A. D. MacKerell Jr., J. Wiorkiewicz-Kuczera, and M. Karplus. An all-atom empirical energy function for the simulation of nucleic acids. *J. Am. Chem. Soc.*, 117(48):11946–11975, 1995.

A. D. MacKerell Jr., D. Bashford, M. Bellott, R. L. Dunbrack Jr., J. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, I. W. E. Reiher, B. Roux, M. Schlenkrich, J. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus. All-hydrogen empirical potential for molecular modeling and dynamics studies of proteins using the CHARMM22 force field. *J. Phys. Chem. B*, 102:3586–3616, 1998.

J. A. McCammon, B. R. Gelin, and M. Karplus. Dynamics of folded proteins. *Nature*, 267:585–590, June 1977.

L. J. McGuffin, K. Bryson, and D. T. Jones. The PSIPRED protein structure prediction server. *Bioinformatics*, 16:404–405, 2000.

N. Metropolis, M. Rosenbluth, A. Rosenbluth, A. Teller, and E. Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21:1087–1092, 1953.

F. Meyer and S. Beucher. Morphological segmentation. *J. Vis Comm. Image R*, 1:21–46, 1990.

W. Minor, J. Steczko, B. Stec, Z. Otwinowski, J. T. Bolin, R. Walter, and B. Axelrod. Crystal structure of soybean lipoxygenase l-1 at 1.4 a resolution. *Biochemistry*, 35:10687–10701, 1996.

M. J. Nelson. Evidence for water coordinated to the active site iron in soybean lipoxygenase-1. *J. Am. Chem. Soc.*, 110:2985–2986, 1988.

M. J. Nelson and S. P. Seitz. The structure and function of lipoxygenase. *Curr. Opin. Struct. Biol.*, 4:878–884, 1994.

P.-O. Norrby and P. Brandt. Deriving force field parameters for coordination complexes. *Coordin. Chem. Rev.*, 212:79–109, 2001.

B. Paizs, J. Baker, S. Suhai, and P. Pulay. Geometry optimization of large biomolecules in redundant internal coordinates. *J. Chem. Phys.*, 113(16): 6566–6572, 2000.

C. Peng, P. Y. Ayala, and H. B. Schlegel. Using redundant internal coordinates to optimize equilibrium geometries and transition states. *J. Comp. Chem.*, 17: 49–56, 1996.

J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten. Scalable molecular dynamics with namd. *J. Comp. Chem.*, 26:1781–1802, 2005.

A. Pidcock, R. E. Richards, and L. M. Venanzi. 195Pt–31P nuclear spin coupling constants and the nature of the trans-effect in platinum complexes. *J. Chem. Soc.*, A:1707–1710, 1966.

N. A. Porter. Mechanisms for the autoxidation of polyunsaturated lipids. *Acc. Chem. Res.*, 19:262–268, 1986.

W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C*. Cambridge University Press, New York, 2 edition, 1992.

P. Pulay and G. Fogarasi. Geometry optimization in redundant internal coordinates. *J. Chem. Phys.*, 96(4):2856–2860, 1992.

P. Pulay, G. Fogarasi, G. Pongor, J. E. Boggs, and Anna Vargha. Combination of theoretical ab initio and experimental information to obtain reliable harmonic force constants. scaled quantum mechanical (sqm) force fields for glyoxal, acrolein, butadiene, formaldehyde, and ethylene. *J. Am. Chem. Soc.*, 105: 7037–7047, 1983.

O. Rådmark, O. Werz, D. Steinhilber, and B. Samuelsson. 5-lipoxygenase: regulation of expression and enzyme activity. *Trends Biochem. Sci.*, 32:332–341, 2007.

S. M. Rapoport and T. Schewe. The maturational breakdown of mitochondria in reticulocytes. *Biochim. Biophys. Acta*, 864:471–495, 1986.

A. K. Rappé, C. J. Casewit, K. S. Colwell, W. A. Goddard III, and W. M. Skiff. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J. Am. Chem. Soc.*, 114:10024–10035, 1992.

G. Rauhut and P. Pulay. Transferable scaling factors for density functional derived vibrational force fields. *J. Phys. Chem.*, 99:3093–3100, 1995.

A. Reed, R. B. Weinstock, and F. Weinhold. Natural population analysis. *J. Chem. Phys.*, 83:735–746, 1985.

A. Reed, L. A. Curtiss, and F. Weinhold. Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint. *Chem. Rev.*, 88:899–926, 1988.

K. W. Rickert and J. P. Klinman. Nature of hydrogen transfer in soybean lipoxygenase 1: Separation of primary and secondary isotope effects. *Biochemistry*, 38:12218–12228, 1999.

S. Riistama, A. Puustinen, A. García-Horsman, S. Iwata, H. Michel, and M. Wikström. Channelling of dioxygen into the respiratory enzyme. *Biochim. Biophys. Acta*, 1275(1–2):1–4, 1996.

S. Riistama, A. Puustinen, M. I. Verkhovsky, J. E. Morgan, H. Michel, and M. Wikstrm. Binding of O_2 and its reduction are both retarded by replacement of valine 279 by isoleucine in cytochrome c oxidase from *Paracoccus Denitrificans*. *Biochemistry*, 39:6365–6372, 2000.

B. Rost and C. Sander. Prediction of protein secondary structure at better than 70% accuracy. *J. Mol. Biol.*, 232:584–599, 1993.

J. Saam, I. Ivanov, M. Walther, H.-G. Holzhütter, and H. Kuhn. Molecular dioxygen enters the active site of 12/15-lipoxygenase via dynamic oxygen access channels. *Proc. Natl. Acad. Sci. USA*, 104:13319–13324, 2007.

R. C. Scarrow, M. G. Trimitsis, C. P. Buck, G. N. Grove, R.A. Cowling, and M. J. Nelson. X-ray spectroscopy of the iron site in soybean lipoxygenase-1: Changes in coordination upon oxidation or addition of methanol. *Biochemistry*, 33:15023–15035, 1994.

J. N. Sharma and L. A. Mohammed. The role of leukotrienes in the pathophysiology of inflammatory disorders: Is there a case for revisiting leukotrienes as therapeutic targets? *Inflammopharmacology*, 14:10–16, 2006.

U. C. Singh and P. A. Kollman. An approach to computing electrostatic charges for molecules. *J. Comp. Chem.*, 5:129, 1984.

F. H. Stillinger and A. Rahman. Improved simulation of liquid water by molecular dynamics. *J. Chem. Phys.*, 60:1545–1557, 1974.

T. R. Stouch and D. E. Williams. Conformational dependence of electrostatic potential-derived charges: Studies of the fitting procedure. *J. Comp. Chem.*, 14:858–866, 1993.

D. Suarez and K. M. Merz Jr. Molecular dynamics simulations of the mononuclear zinc-β-lactamase from *Bacillus Cereus*. *J. Am. Chem. Soc.*, 123:3759–3770, 2001.

R. F. Tilton, I. D. Kuntz, and G.A. Petsko. Cavities in proteins: Structure of a metmyoglobin-xenon complex solved to 1.9 å. *Biol. Cyber.*, 23:2849–2857, 1984.

W. F. van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. R. P. Scott, and I. G. Tironi. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*. vdf Hochschulverlag AG an der ETH Zürich and BIOMOS b.v., Zürich, Groningen, 1996.

K. van Leyen, R. M. Duvoisin, H. Engelhardt, and M. Wiedmann. A function for lipoxygenase in programmed organelle degradation. *Nature*, 395:392–395, 1998.

A. Vedani and D. W. Huhta. A new force field for metalloproteins. *J. Am. Chem. Soc.*, 112:4759–4767, 1990.

M. Walther, M. Anton, M. Wiedmann, R. Fletterick, and H. Kuhn. The n-terminal domain of the reticulocyte-type 15-lipoxygenase is not essential for enzymatic activity but contains determinants for membrane binding. *J. Biol. Chem.*, 277:27360–27366, 2002.

J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case. Development and testing of a general amber force field. *J. Comp. Chem.*, 25: 1157–1174, 2004.

E. B. Wilson, J. C. Decius, and P. C. Cross. *Molecular Vibrations*. Dover Publications, New York, 1955.

S. Yamamoto. Mammalian lipoxygenases: molecular and catalytic properties. *Prostaglandins Leukot. Essent. Fatty Acids*, 35:219–229, 1989.

L. Zhang and J. Hermans. Hydrophilicity of cavities in proteins. *Proteins: Struct., Func., Gen.*, 24:433–438, 1996.

Y. Zhang, Q-F. Gan, E. G. Pavel, E. Sigal, and E. I. Solomon. Epr definition of the non-heme ferric active sites of mammalian 15-lipoxygenase: Major spectral difference relative to huan 5-lipoxygenases and plant lipoxygenases and their ligand field origin. *J. Am. Chem. Soc.*, 117:7422–7427, 1995.

Appendix A

Force field parameter determination

The parameterization schemes for CHARMM and the other force fields mentioned above are quite similar with regard to the bonded interactions, while greater differences can be found in the derivation of charges (see below). Force fields are parameterized against a combination of experimental and *ab initio* target data, i.e. properties of a molecule that the force field should be able to reproduce. Examples are equilibrium geometries, vibrational frequencies, dipole moments, molecular volume, or thermodynamic properties like heats of vaporization or free energies of solvation. Equilibrium geometries are typically obtained from X-ray crystal structures or electron diffraction data. Force constants for bonded interactions can be determined using vibrational frequencies from infrared and Raman spectroscopy. Where reliable experimental target data are missing, e.g. for many organic compounds, those can be obtained from quantum mechanical calculations.

Transferable parameter sets

In order to limit the effort and to make force field parameters transferable to other chemically equivalent molecular moieties, model compounds representing basic molecular building blocks, functional groups (e.g. benzene ring or acetate) or linkages (e.g. ester bond or peptide bond) are parameterized. These building blocks can than be combined into bigger molecules like amino acids or lipids or proteins. Atoms are categorized in types according to their chemical environment. There exist, for example, atom types for oxygen in carboxylate, carbonyl and hydroxyl groups. These atom types allow the definition of reusable parameters for different bonds, angles, dihedrals or vdW interaction pairs.

Charges

Molecular charges are composed of nuclear charges and electrons, but due to the delocalized nature of the electronic wavefunctions, especially for valence electrons that form bonds, charge is not an atomic property. In classical force fields the continuous charge distribution is mapped onto point charges located at the atomic nuclei.

The charges for CHARMM are chosen such that the force field based interaction distances and energies of TIP3P water models [Jorgensen et al., 1983] with model compounds reproduce scaled *ab initio* 6-31G* interaction distances and energies between water and model compounds [MacKerell Jr. et al., 1995]. For polar species, the 6-31G* interaction energies are scaled by a factor of 1.16, which compensates for the lack of explicit polarization effects in the potential energy functions of the force field. No scaling is applied for charged species, because in this case HF/6-31G* interaction energies yield charge distributions giving satisfactory agreement with heats and free energies of solvation. Further, the quantum chemical dipole moments of the model compounds should be reflected in the point charge model. The parameterization of charges for OPLS and GROMOS follows a similar scheme whereas with TIP4P [Jorgensen et al., 1983] and SPC [Berendsen et al., 1981] different water models are used.

For AMBER a completely different approach has been chosen. In restricted electrostatic potential fitting (RESP) [Cornell et al., 1993] atom centered point charges are adjusted such as to reproduce the *ab initio* molecular electrostatic potential for a large number test points on several shells around the molecule. In the two stage fitting procedure properties like total charge or symmetry of methyl hydrogen atoms, are conserved and certain restraints are applied to buried atoms in order to compensate bad fitting due to statistically underrepresented buried atoms.

Appendix B

Hessian of the equilibrium potential

Background and details of the following are described e.g. in [Wilson et al., 1955]. The parameters of bonded interactions in classical force fields are based on a local harmonic approximation of the potential energy at equilibrium. We can express the potential energy V around the equilibrium in form of a Taylor expansion in the cartesian displacement coordinates x_i:

$$\begin{aligned} V &= V_0 + \sum_{i=1}^{3N} \left(\frac{\partial V}{\partial x_i}\right)_0 x_i + \frac{1}{2} \sum_{i,j=1}^{3N} \left(\frac{\partial^2 V}{\partial x_i \partial x_j}\right)_0 x_i x_j + \dots \\ &= V_0 + \sum_{i=1}^{3N} g_i x_i + \frac{1}{2} \sum_{i,j=1}^{3N} H_{ij} x_i x_j + \dots \end{aligned} \tag{B.1}$$

We eliminate V_0 by defining the origin of the energy such that $V_0 = 0$. Further, the equilibrium geometry denotes the minimum of the potential where the first derivatives also vanish:

$$\left(\frac{\partial V}{\partial x_i}\right)_0 = g_i = 0, \qquad i = 1, 2, \dots, 3N \tag{B.2}$$

Thus, for small amplitudes of vibrations where higher terms can be neglected we obtain the following local approximation for the potential energy:

$$V = \frac{1}{2} \sum_{i,j=1}^{3N} \left(\frac{\partial^2 V}{\partial x_i \partial x_j}\right)_0 x_i x_j = \sum_{i,j=1}^{3N} H_{ij} x_i x_j \tag{B.3}$$

The second derivatives of the potential energy with respect to the coordinates, H_{ij}, form a matrix **H**, also called the cartesian Hessian matrix.

Coordinate transformation

The cartesian Hessian can be transformed into redundant internal coordinates consisting of the bonds, bond angles, and dihedrals that are used in the force field. The diagonal elements of this internal Hessian matrix **F** contain the force constants for the corresponding internal coordinates, while the off-diagonals denote the couplings between the different modes. The transformation from cartesian to redundant internal coordinates is described in [Pulay and Fogarasi, 1992, Paizs et al., 2000, Peng et al., 1996] and will be summarized in the following.

The potential energy V can, instead of cartesian coordinates, also be expanded with respect to internal coordinates $\mathbf{q} = (q_1, q_2, \ldots, q_M)$:

$$\begin{aligned} V &= V_0 + \sum_{i=1}^{3M} \left(\frac{\partial V}{\partial q_i}\right)_0 q_i + \frac{1}{2} \sum_{i,j=1}^{3M} \left(\frac{\partial^2 V}{\partial q_i \partial q_j}\right)_0 q_i q_j + \ldots \\ &= V_0 + \sum_{i=1}^{3M} f_i q_i + \frac{1}{2} \sum_{i,j=1}^{3M} F_{ij} q_i q_j + \ldots \end{aligned} \tag{B.4}$$

Further, the internal coordinates can be expanded in a power series with respect to the cartesian displacement coordinates $\mathbf{x} = (x_1, x_2, \ldots, x_M)$

$$q_i = \mathbf{B}_i \mathbf{x} + \frac{1}{2} \mathbf{x}^+ \mathbf{C}_i \mathbf{x} + \ldots \tag{B.5}$$

while $\mathbf{B}_i$ is the i^{th} row of Wilson's vibrational matrix. Detailed documentation how to construct this matrix from the molecular geometry can be found in Wilson et al. [1955]. Combining equations B.1 and B.4, one obtains (in a more convenient matrix notation)

$$V = V_0 + \mathbf{g}^+ \mathbf{x} + \frac{1}{2} \mathbf{x}^+ \mathbf{H} \mathbf{x} + \cdots = V_0 + \mathbf{f}^+ \mathbf{q} + \frac{1}{2} \mathbf{q}^+ \mathbf{F} \mathbf{q} + \ldots , \tag{B.6}$$

where the row vectors $\mathbf{g}$ and $\mathbf{f}$ are the gradients in cartesian and internal coordinates respectively. Matrices $\mathbf{H}$ and $\mathbf{F}$ denote the corresponding Hessians. Substituting expression (B.5) in (B.6) yields

$$\begin{aligned} V = V_0 &+ \sum_{i=1}^{3M} f_i \mathbf{B}_i \mathbf{x} + \frac{1}{2} \sum_{i=1}^{3M} f_i \mathbf{x} \mathbf{C}_i \mathbf{x} \\ &+ \frac{1}{2} \sum_{i,j=1}^{3M} F_{ij} \left(\mathbf{B}_i \mathbf{x} + \frac{1}{2} \mathbf{x} \mathbf{C}_i \mathbf{x}\right) \left(\mathbf{B}_j \mathbf{x} + \frac{1}{2} \mathbf{x} \mathbf{C}_j \mathbf{x}\right) . \end{aligned} \tag{B.7}$$

Our goal is to equate like powers in (B.7) and the left side of (B.6), in order to obtain expressions relating gradients, as well as Hessians, in cartesian and internal coordinates. We are only interested in the second order approximation, hence, we can omit all higher order terms that result from the multiplication of the two brackets in the right most term and get

$$V = V_0 + \mathbf{f}^{+}\mathbf{B}\mathbf{x} + \frac{1}{2}\mathbf{x}^{+}\sum_{i=1}^{3M} f_i \mathbf{C}_i \mathbf{x} + \sum_{i=1}^{3M} \mathbf{x}^{+} F_{ij} \mathbf{B}_i \mathbf{B}_j \mathbf{x}. \tag{B.8}$$

Now, one can easily identify the following relations that express the cartesian gradient vector $\mathbf{g}$ and Hessian $\mathbf{H}$ in terms of their counterpart in internal coordinates.

$$\mathbf{g} = \mathbf{B}^{+}\mathbf{f} \tag{B.9}$$

$$\mathbf{H} = \sum_{i=1}^{3M} \mathbf{f}_i \mathbf{C}_i + \mathbf{B}^{+}\mathbf{F}\mathbf{B} \tag{B.10}$$

Since we are evaluating force constants for equilibrium geometry, the forces (negative gradients of the energy) are zero, and (B.10) simplifies to

$$\mathbf{H} = \mathbf{B}^{+}\mathbf{F}\mathbf{B} \tag{B.11}$$

$\mathbf{B}$ is a $M \times (3N)$ matrix, and thus possesses no inverse. However, a right inverse can be defined as

$$\mathbf{A} = \mathbf{u}\mathbf{B}^{+}(\mathbf{B}\mathbf{u}\mathbf{B}^{+})^{-1}, \tag{B.12}$$

such that

$$\mathbf{B}\mathbf{A} = \mathbf{I}, \tag{B.13}$$

where $\mathbf{u}$ is a $3N \times 3N$ nonsingular matrix, typically the unit matrix. Thus, after introducing $\mathbf{G} = \mathbf{B}\mathbf{u}\mathbf{B}^{+}$, we get

$$\mathbf{f} = \mathbf{A}^{+}\mathbf{f} = \mathbf{G}^{-1}\mathbf{B}\mathbf{u}\mathbf{g} \tag{B.14}$$

$$\mathbf{F} = \mathbf{A}^{+}\mathbf{F}\mathbf{A} = \mathbf{G}^{-1}\mathbf{B}\mathbf{u}\mathbf{H}\mathbf{u}\mathbf{B}^{+}\mathbf{G}^{-1}. \tag{B.15}$$

This expression could, in principle be used to convert the cartesian Hessian $\mathbf{H}$ into its internal coordinate form $\mathbf{F}$. But if we are using a redundant internal coordinate set, $\mathbf{G}$ is singular, which means that is has no inverse. Pulay and Fogarasi [1992] suggested using the generalized inverse $\mathbf{G}^{-}$. The latter is denoted by a minus sign superscript, to distinguish it from the ordinary inverse. It can be obtained by diagonalizing $\mathbf{G}$

$$\mathbf{U}^{+}\mathbf{G}\mathbf{U} = \begin{bmatrix} \mathbf{\Lambda} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \tag{B.16}$$

inverting the nonzero eigenvalues $\mathbf{\Lambda}$, and transforming it back

$$\mathbf{G}^{-} = \mathbf{U} \left[\begin{array}{cc} \mathbf{\Lambda}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{array} \right] \mathbf{U}^{+}. \tag{B.17}$$

The eigenvalues of singular matrices can be computed using singular value decomposition [see e.g. Press et al., 1992]. Finally, the equation for the Hessian in redundant internal coordinates becomes

$$\mathbf{F} = \mathbf{G}^{-}\mathbf{BuHuB}^{+}\mathbf{G}^{-}. \tag{B.18}$$

Appendix C

Enzyme characteristics

To verify that site-directed mutagenesis did not completely alter the functional enzyme characteristics, the relevant mutants were characterized with respect to their basic enzymatic properties such as reaction kinetics and product specificity. In Fig. C.1 the arachidonic acid oxygenation products formed by the different mutants are shown. It can be seen that the L367F, L367W, L367K, and L367E mutants exhibit a similar positional specificity as the wild-type enzyme. Next, we determined basic kinetic constants for the reaction of the relevant mutants with linoleic acid (Table 4.4). Although the mutants exhibited a variable specific activity they were all catalytically active. The most important mutant for the oxygen binding studies, L367F, exhibited a similar catalytic efficiency as the wild type enzyme. Taken together, our experimental data demonstrated that the mutagenesis strategy did not severely disturb the catalytic function.

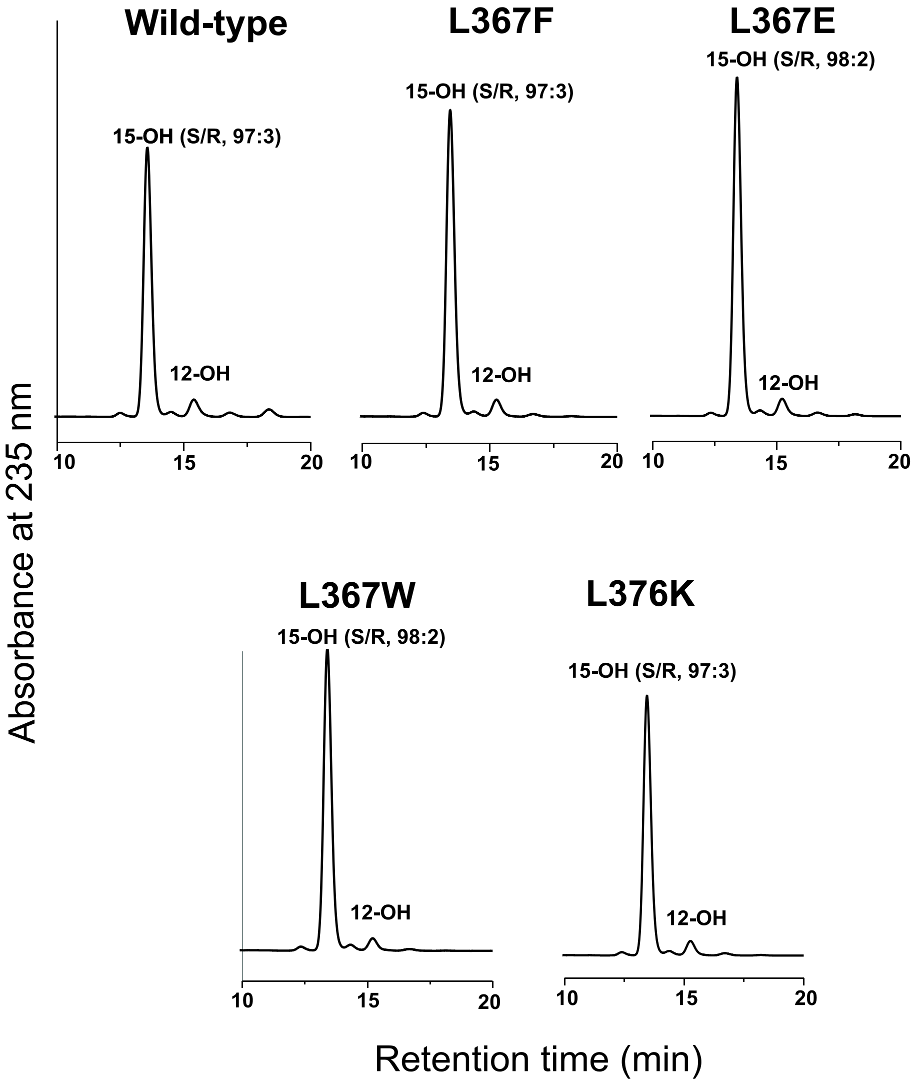


Figure C.1: HPLC analysis of arachidonic acid oxygenation products formed by different 12/15-LOX mutants. The 12/15-LOX mutants were expressed as his-tag fusion proteins in E. coli and purified from the lysate supernatants by affinity chromatography on a Ni-agarose column. Aliquots of the final enzyme preparations were incubated in 0.1 M phosphate buffer, pH 7.4 for 5 min with 100 μM arachidonic acid. The oxygenation products formed were reduced with sodium borohydride and analyzed by RP-HPLC. The stereochemistry of the main product 15-HETE was determined by SP-HPLC and the S/R ratio is given.

Appendix D

Kinetic Model for the LOX-Reaction

Kinetic considerations in this work are based on reaction scheme 3.2 which is listed here again

$$\mathrm{E} \xrightleftharpoons[k_{-1}]{k_{+1}[\mathrm{S}]} \mathrm{ES} \xrightarrow{k_{\mathrm{S}}} \mathrm{ES}^{*} \xrightleftharpoons[k_{-0}]{k_{+0}[\mathrm{O_2}]} \mathrm{ES^{*}O_2} \xrightarrow{k_{\mathrm{P}}} \mathrm{P}$$

and the according kinetic equations

$$\frac{\mathrm{d[ES]}}{\mathrm{d}t} = k_{+1}[\mathrm{E}][\mathrm{S}] - (k_{-1} + k_{\mathrm{S}})[\mathrm{ES}] \tag{D.1}$$

$$\frac{\mathrm{d[ES^{*}]}}{\mathrm{d}t} = k_{\mathrm{S}}[\mathrm{ES}] - k_{+0}[\mathrm{ES^{*}}][\mathrm{O_2}] \tag{D.2}$$

$$\frac{\mathrm{d[ES^{*}O_2]}}{\mathrm{d}t} = k_{+0}[\mathrm{ES^{*}}][\mathrm{O_2}] - k_{-0}[\mathrm{ES^{*}O_2}] - k_{\mathrm{P}}[\mathrm{ES^{*}O_2}] \tag{D.3}$$

$$\frac{\mathrm{d[P]}}{\mathrm{d}t} = k_{\mathrm{P}}[\mathrm{ES^{*}O_2}] = v \tag{D.4}$$

where the notation listed below is used for the concentrations and rate constants:

$[\mathrm{O_2}]$ – external dioxygen
$[\mathrm{S}]$ – fatty acid substrate
$[\mathrm{ES^{*}}]$ – enzyme fatty acid radical complex
$[\mathrm{ES^{*}O_2}]$ – enzyme fatty acid radical complex with channelled oxygen
$[\mathrm{P}]$ – reaction product (hydroperoxy fatty acid)

k_{+1} – on-rate constant for binding of fatty acid
k_{-1} – off-rate constant for release of fatty acid
k_{S} – rate constant for hydrogen abstraction
k_{P} – rate constant for product formation
k_{+0} – on-rate constant for oxygen uptake
k_{-0} – off-rate constant for oxygen release

In the following the derivation of the stationary rate law (3.3) and the ratio $k_{\mathrm{cat}}/K_{\mathrm{M}}^{O_2}$ (3.4) from the Methods section will be shown in detail.

Since the concentration of the enzyme is small with respect to the substrate constant concentrations of the intermediates, i.e. $\frac{\mathrm{d[ES]}}{\mathrm{d}t} = 0$, $\frac{\mathrm{d}ES^*}{\mathrm{d}t} = 0$ and $\frac{\mathrm{d[ES^*O_2]}}{\mathrm{d}t} = 0$ can be assumed (quasi stationarity). Hence we obtain

$$[\mathrm{ES}] = \frac{[\mathrm{E}][\mathrm{S}]}{K_{\mathrm{M}}^{\mathrm{S}}} \tag{D.5}$$

$$[\mathrm{ES}^*] = \frac{k_{\mathrm{S}}[\mathrm{ES}] + k_{-0}[\mathrm{ES^*O_2}]}{k_{+0}[\mathrm{O_2}]} \tag{D.6}$$

$$[\mathrm{ES^*O_2}] = \frac{[\mathrm{ES}^*][\mathrm{O_2}]}{K^{\sharp}} \tag{D.7}$$

with the apparent Michaelis constant for the fatty acid substrate

$$K_{\mathrm{M}}^{\mathrm{S}} = \frac{k_{-1} + k_{\mathrm{S}}}{k_{+1}} \tag{D.8}$$

and

$$K^{\sharp} = \frac{k_{-0} + k_{\mathrm{P}}}{k_{+0}}. \tag{D.9}$$

Inserting expression D.7 into D.6 yields

$$[\mathrm{ES}^*] = \frac{k_{\mathrm{S}}}{k_{\mathrm{P}}} \frac{K^{\sharp}[\mathrm{ES}]}{[\mathrm{O_2}]} \tag{D.10}$$

and therefore

$$[\mathrm{ES^*O_2}] = \frac{k_{\mathrm{S}}}{k_{\mathrm{P}}}[\mathrm{ES}]. \tag{D.11}$$

Conservation of the total enzyme concentration $[\mathrm{E_{tot}}]$ demands that

$$[\mathrm{E_{tot}}] = [\mathrm{E}] + [\mathrm{ES}] + [\mathrm{ES}^*] + [\mathrm{ES^*O_2}]. \tag{D.12}$$

Putting D.5, D.10 and D.11 into the conservation condition D.12 we get

$$[\mathrm{E}] = \frac{[\mathrm{E_{tot}}]}{1 + \frac{[\mathrm{S}]}{K_{\mathrm{M}}^{\mathrm{S}}}\left(1 + \frac{k_{\mathrm{S}}}{k_{\mathrm{P}}}\left(1 + \frac{K^{\sharp}}{[\mathrm{O_2}]}\right)\right)} \tag{D.13}$$

for the concentration of substrate free enzyme. The reaction rate thus reads

$$v = k_{\mathrm{P}}[\mathrm{ES^*O_2}] = k_{\mathrm{S}}[\mathrm{ES}] = \frac{k_{\mathrm{S}}\frac{[\mathrm{S}]}{K_{\mathrm{M}}^{\mathrm{S}}}[\mathrm{E_{tot}}]}{1 + \frac{[\mathrm{S}]}{K_{\mathrm{M}}^{\mathrm{S}}}\left(1 + \frac{k_{\mathrm{S}}}{k_{\mathrm{P}}}\left(1 + \frac{K^{\sharp}}{[\mathrm{O_2}]}\right)\right)} \tag{D.14}$$

Assuming substrate saturation (i.e. $[\mathrm{S}] \gg K_\mathrm{M}^\mathrm{S}$) the rate equation D.14 simplifies to

$$v = \frac{k_\mathrm{cat}[\mathrm{O_2}][\mathrm{E_{tot}}]}{K_\mathrm{M}^\mathrm{O_2} + [\mathrm{O_2}]} \tag{D.15}$$

with

$$k_\mathrm{cat} = \frac{k_\mathrm{S}}{1 + \frac{k_\mathrm{S}}{k_\mathrm{P}}} \tag{D.16}$$

and

$$K_\mathrm{M}^\mathrm{O_2} = \frac{\frac{1}{K_0}\frac{k_\mathrm{S}}{k_\mathrm{P}} + \frac{k_\mathrm{S}}{k_{+0}}}{1 + \frac{k_\mathrm{S}}{k_\mathrm{P}}}. \tag{D.17}$$

Here $K_0 = k_{+0}/k_{-0}$ is the equilibrium constant for oxygen diffusion through the channels. As several studies have revealed hydrogen abstraction to be the rate limiting step in LOX catalysis it holds that $k_\mathrm{S} \ll k_\mathrm{P}$. Thus the catalytic constant (turnover number)

$$k_\mathrm{cat} = k_\mathrm{S} \tag{D.18}$$

is identical with the rate of hydrogen abstraction and the apparent Michaelis constant

$$K_\mathrm{M}^\mathrm{O_2} = \frac{1}{K_0}\frac{k_\mathrm{S}}{k_\mathrm{P}} + \frac{k_\mathrm{S}}{k_{+0}}. \tag{D.19}$$

depends on the rate constants for oxygen exchange as well as the rate constants for hydrogen abstraction and oxygen insertion.

From relations D.18 and D.19 we finally get to expression (3.4) from the Methods:

$$\frac{k_\mathrm{cat}}{K_\mathrm{M}^\mathrm{O_2}} = \frac{1}{\frac{1}{K_0 k_\mathrm{P}} + \frac{1}{k_{+0}}}. \tag{D.20}$$

Lowered values of this ratio as observed for the channel mutants may thus be due to lowered values of k_{+0} or $K_0 k_\mathrm{P}$. We assume, however, no significant change of the association constant K_0 for oxygen uptake (the calculated oxygen occupation probabilities at the active site are very similar in wild type and mutants). Further, we exclude a significant alteration of the rate constant k_P. Combining under these premises equation D.20 for wild type (wt) and the mutant (mut)

$$\frac{k_{+0}^\mathrm{wt}}{k_{+0}^\mathrm{mut}} = 1 + k_{+0}^\mathrm{wt}\left(\left(\frac{K_M^\mathrm{O_2}}{k_\mathrm{cat}}\right)^\mathrm{mut} - \left(\frac{K_M^\mathrm{O_2}}{k_\mathrm{cat}}\right)^\mathrm{wt}\right) \tag{D.21}$$

we see that the observed changes in the ratio $\frac{k_\mathrm{cat}}{K_\mathrm{M}^\mathrm{O_2}}$ can be accounted for by changes in k_{+0}.

Appendix E

Coordination complex properties

The DFT geometry together with the vibrational force constants reveals a number of chemical features of the complex. The coordinative bonds between iron and the ligands are constituted by electrostatic interactions and a covalent part due to the donation of electrons from the ligands into the singly occupied iron d-orbitals. Table 4.1 lists the bond lengths and force constants of all bonds between iron center and ligands. It is obvious that the charged ligands acetate and hydroxide, being excellent electron donors, are bonded much tighter to the metal center than neutral ligands. The ligand–metal bonds in the Fe^{3+} complex exhibit a remarkable trans-influence [Pidcock et al., 1966], i.e. σ-bonds with strong covalent character connecting a ligand and the metal center weaken the bond of the trans-oriented ligand. This can be observed for ligands His361 and His541, located trans to the charged acetate and hydroxide groups, that show significantly decreased force constants compared to the other two histidines.

Natural bond orbital (NBO) analysis [Reed et al., 1988], which attempts to project molecular orbitals on orbitals that define Lewis bonds, helps rationalizing this fact. Consider for example the OH^--ligand interacting in two ways with iron. Figure E.1A shows that an oxygen sp_2-hybridized orbital forms a σ-bond to a formally empty d-orbital of iron. At the same time, the opposite His541 uses the same d-orbital for its bond to the center involving a nitrogen sp_2-type orbital. Being the stronger electron donor, OH^- forms a tighter bond to the center and claims the d-orbital more vigorously than the imidazole. Thus the bond with imidazole is weakened. Of course, the final bond strength is due to a combination with other effects for example the same d-orbital also interacts with acetate and His361. Further, electrons are donated from the doubly occupied p-type lone-pair orbital of the hydroxide ion into another singly occupied d-orbital at iron, forming a coordinative pi-bond (fig. E.1B).

The main difference between the ferric and the ferrous form is that hydroxide is bound 0.4 Å closer to iron than water due to the electrostatic attraction between the hydroxy anion and iron. As the formerly charged hydroxide becomes

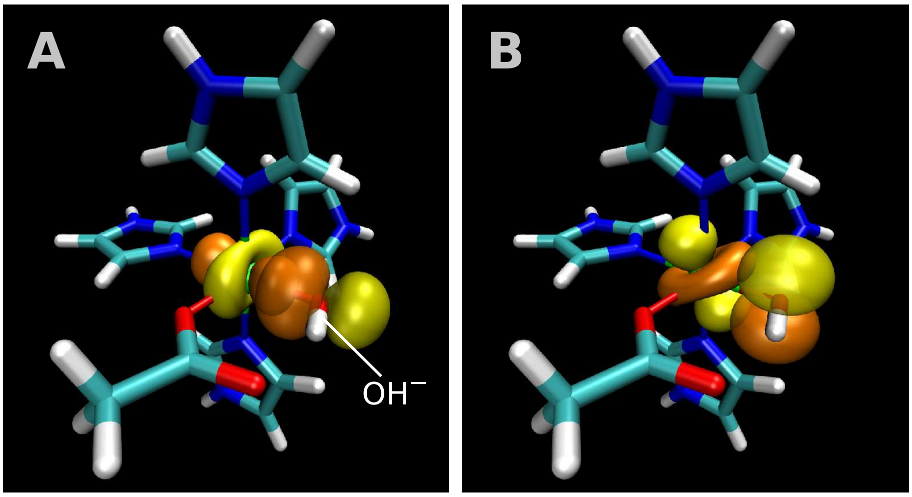


Figure E.1: Coordinative bonds of iron(III) d-orbitals with ligands. (A) σ-bond with sp^3-type orbital of hydroxide. **(B)** π-bond with p-orbital of hydroxide.

neutral water in the Fe^{2+} complex the trans influence on His541 vanishes, which might be a reason for the shortening of the bond an the weaker force constant. The reduced electrostatic attraction between the Ile663 carboxylate group and Fe^{2+} could explain that this bond becomes weaker upon reduction of the complex.

Abbreviations

abbreviation	explanation
B3LYP	Becke's three parameter hybrid functional using the Lee-Yang-Parr correlation functional
CHARMM	chemistry at Harvard molecular mechanics
COX	cyclooxygenase
DNA	desoxyribonucleic acid
DFT	density functional theory
ESP	electrostatic potential fitting
EXAFS	extended x-ray absorption fine structure
GAFF	generalized AMBER force field
GROMOS	Groningen molecular simulation
HF	Hartree-Fock
HPETE	hydroperoxy-eicosatetraene
ILS	implicit ligand sampling
LOX	lipoxygenase
MD	molecular dynamics
MMFF	Merck molecular force field
NBO	natural bond orbital
NPA	natural population analysis
NVE	ensemble with constant particle number N, volume V and temperature T
OPLS	optimized potentials for liquid simulations
RESP	restricted electrostatic potential fitting
SPC	simple point charge
TIP3P	transferable intermolecular potential 3 point
UFF	universal force field
vdW	van der Waals

Danksagung

Prof. Holzhütter möchte ich danken für das große Vertrauen, dass er in mich gesetzt hat, und für die gewährte Freiheit meine eigenen Wege zu gehen bedanken.

Prof. Hartmut Kühn, Dr. Igor Ivanov und Dr. Matthias Walther danke ich für die jahrelange gute Zusammenarbeit und das große Vertrauen in den Erfolg meiner Bemühungen.

Dr. Bernd Kallies danke ich für die Einführung in die Quantenchemie und die unermüdliche Hilfe mit der entsprechenden Software.

Dr. Grunenberg bin ich sehr dankbar für zahlreiche fruchtbringende Diskussionen über Kraftfelder.

Prof. Klaus Schulten und Prof. Zaida Luthey-Schulten gebühren Dank für die Ermutigung und Unterstützung bei der Entwicklung von PARATOOL.

John Stone für die Berücksichtigung vieler meiner Sonderwünsche bei der Entwicklung von VMD.

Maren, ich danke Dir sehr herzlich für all die Unterstützung und Hilfe, die Du mir zu Teil werden ließest.

Selbständigkeitserklärung

Ich erkläre, dass ich diese Arbeit selbständig und nur mit den angegebenen Hilfmitteln angefertigt habe und daß alle Stellen, die dem Wortlaut oder dem Sinne nach anderen Werken entnommen sind, durch Angaben der Quellen als Entlehnung kenntlich gemacht worden sind.

Ich erkläre die Kenntnisnahme der dem Verfahren zugrunde liegenden Promotionsordnung der Mathematisch-Naturwissenschaftlichen Fakultät I der Humboldt-Universität zu Berlin.

Ich besitze noch keinen Doktorgrad und habe mich nicht anderwärts um einen solchen beworben.